PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



(51) International Patent Classification 7:		(1) International Publication Number: WO 00/393
C12N 15/49, C07K 14/16	A2	13) International Publication Date: 6 July 2000 (06.07)
(21) International Application Number: PCT/USS (22) International Filing Date: 30 December 1999 (3 (30) Priority Data: 60/114,495 60/156,670 29 September 1998 (31.12.98 60/156,670 29 September 1999 (29.09.99) (71) Applicant: CHIRON CORPORATION [US/US]; 4560 Street, Emeryville, CA 94608 (US). (72) Inventors: BARNETT, Susan; Chiron Corporation, 45 ton Street – R440, Emeryville, CA 94608 (US). H Karin; Chiron Corporation, 4560 Horton Street Emeryville, CA 94608 (US). MARTIN, Eric; Chir poration, 4560 Horton Street – R440, Emeryville, C (US). (74) Agents: DOLLARD, Anne, S.; Chiron Corporation, tual Property – R440, P.O. Box 8097, Emeryv 94662–8097 (US) et al.	3) U 9) U 0 Horto 560 Ho ARTO – R44 ron Co 'A 9460	(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZA, ZARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, NRU, TJ, TM), European patent (AT, BE, CH, CY, DE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), Opatent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, NNE, SN, TD, TG). Published Without international search report and to be republis upon receipt of that report.
(54) Title: MODIFIED HIV ENV POLYPEPTIDES (57) Abstract Polynucleotide encoding modified HIV Env polypeptic of the CD4 binding region. Methods of diagnosis, treatment	des are	sclosed. The Env polypeptides are modified so as to expose at least vention using the polynucleotides and polypeptides are also provides

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

		_					
AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	$\mathbf{T}\mathbf{T}$	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

MODIFIED HIV ENV POLYPEPTIDES

Technical Field

5

10

15

20

25

30

The invention relates generally to modified HIV envelope (Env) polypeptides which are useful as immunizing agents or for generating an immune response in a subject, for example a cellular immune response or a protective immune response. More particularly, the invention relates Env polypeptides such as gp120, gp140 or gp160, wherein at least one of the native β -sheet configurations has been modified. The invention also pertains to methods of using these polypeptides to elicit an immune response against a broad range of HIV subtypes.

Background of the Invention

The human immunodeficiency virus (HIV-1, also referred to as HTLV-III, LAV or HTLV-III/LAV) is the etiological agent of the acquired immune deficiency syndrome (AIDS) and related disorders. (see, e.g., Barre-Sinoussi, et al., (1983) Science 220:868-871; Gallo et al. (1984) Science 224:500-503; Levy et al., (1984) Science 225:840-842; Siegal et al., (1981) N. Engl. J. Med. 305:1439-1444). AIDS patients usually have a long asymptomatic period followed by the progressive degeneration of the immune system and the central nervous system. Replication of the virus is highly regulated, and both latent and lytic infection of the CD4 positive helper subset of T-lymphocytes occur in tissue culture (Zagury et al., (1986) Science 231:850-853). Molecular studies of HIV-1 show that it encodes a number of genes (Ratner et al., (1985) Nature 313:277-284; Sanchez-Pescador et al., (1985) Science 227:484-492), including three structural genes -- gag, pol and env -- that are common to all retroviruses. Nucleotide sequences from viral genomes of other retroviruses, particularly HIV-2 and simian immunodeficiency viruses, SIV (previously referred to as STLV-III), also contain these structural genes. (Guyader et al., (1987) Nature 326:662-669; Chakrabarti et al., (1987) Nature

The envelope protein of HIV-1, HIV-2 and SIV is a glycoprotein of about 160 kd (gp160). During virus infection of the host cell, gp160 is cleaved by host cell proteases to form gp120 and the integral membrane protein, gp41. The gp41 portion is anchored in the

5

10

15

20

25

30

membrane bilayer of virion, while the gp120 segment protrudes into the surrounding environment. gp120 and gp41 are more covalently associated and free gp120 can be released from the surface of virions and infected cells.

As depicted in Figure 1, crystallography studies of the gp120 core polypeptide indicate that this polypeptide is folded into two major domains having certain emanating structures. The inner domain (inner with respect to the N and C terminus) features a two-helix, two-stranded bundle with a small five-stranded β -sandwich at its termini-proximal end and a projection at the distal end from which the V1/V2 stem emanates. The outer domain is a staked double barrel that lies along side the inner domain so that the outer barrel and inner bundle axes are approximately parallel. Between the distal inner domain and the distal outer domain is a four-stranded bridging sheet which holds a peculiar minidomain in contact with, but distinct from, the inner, the outer domain, and the V1/V2 domain. The bridging sheet is composed of four β -strand structures (β -3, β -2, β -21, β -20, shown in Figure 1). The bridging region can be seen in Figure 1 packing primarily over the inner domain, although some surface residues of the outer domain, such as Phe 382, reach into the bridging sheet to form part of its hydrophobic core.

The basic unit of the β -sheet conformation of the bridging sheet region is the β -strand which exists as a less tightly coiled helix, with 2.0 residues per turn. The β -strand conformation is only stable when incorporated into a β -sheet, where hydrogen bonds with close to optimal geometry are formed between the peptide groups on adjacent β -strands; the dipole moments of the strands are also aligned favorably. Side chains from adjacent residues of the same strand protrude from opposite sides of the sheet and do not interact with each other, but have significant interactions with their backbone and with the side chains of neighboring strands. For a general description of β -sheets, see, e.g., T.E. Creighton, <u>Proteins: Structures and Molecular Properties</u> (W.H. Freeman and Company, 1993); and A.L. Lehninger, Biochemistry (Worth Publishers, Inc., 1975).

The gp120 polypeptide is instrumental in mediating entry into the host cell. Recent studies have indicated that binding of CD4 to gp120 induces a conformational change in Env that allows for binding to a co-receptor (*e.g.*, a chemokine receptor) and subsequent entry of the virus into the cell. (Wyatt, R., et al. (1998) *Nature* 393:705-711; Kwong, P., et al. (1998) *Nature* 393:648-659). Referring again to Figure 1, CD4 is bound into a depression formed at the interface of the outer domain, the inner domain and the bridging sheet of gp120.

Immunogenicity of the gp120 polypeptide has also been studied. For example, individuals infected by HIV-1 usually develop antibodies that can neutralize the virus in *in vitro* assays, and this response is directed primarily against linear neutralizing determinants in the third variable loop of gp120 glycoprotein (Javaherian, K., et al. (1989) *Proc. Natl. Acad. Sci.* 86:6786-6772; Matsushita, M., et al. (1988) *J. Virol.* 62:2107-2144; Putney, S., et al. (1986) *Science* 234:1392-1395; Rushe, J. R., et al. (1988) *Proc. Nat. Acad. Sci. USA* 85: 3198-3202.). However, these antibodies generally exhibit the ability to neutralize only a limited number of HIV-1 strains (Matthews, T. (1986) *Proc. Natl. Acad. Sci. USA*. 83:9709-9713; Nara, P. L., et al. (1988) *J. Virol.* 62:2622-2628; Palker, T. J., et al. (1988) *Proc. Natl. Acad. Sci. USA*. 85:1932-1936). Later in the course of HIV infection in humans, antibodies capable of neutralizing a wider range of HIV-1 isolates appear (Barre-Sinoussi, F., et al. (1983) *Science* 220:868-871; Robert-Guroff, M., et al. (1985) *Nature* (London) 316:72-74; Wcis, R., et al. (1985) *Nature* (London) 316:69-72; Weis, R., et al. (1986) *Nature* (London) 324:572-575).

Recent work done by Stamatatos et al (1998) *AIDS Res Hum Retroviruses*14(13):1129-39, shows that a deletion of the variable region 2 from a HIV-1_{SF162} virus, which utilizes the CCR-5 co-receptor for virus entry, rendered the virus highly susceptible to serum-mediated neutralization. This V2 deleted virus was also neutralized by sera obtained from patients infected not only with clade B HIV-1 isolates but also with clade A, C, D and F HIV-1 isolates. However, deletion of the variable region 1 had no effect. Deletion of the variable regions 1 and 2 from a LAI isolate HIV-I_{IIIB} also increased the susceptibility to neutralization by monoclonal antibodies whose epitopes are located within the V3 loop, the CD4-binding site, and conserved gp120 regions (Wyatt, R., et al. (1995) *J Virol.* 69:5723-5733). Rabbit immunogenicity studies done with the HIV-1 virus with deletions in the V1/V2 and V3 region from the LAI strain, which uses the CXCR4 co-receptor for virus entry, showed no improvement in the ability of Env to raise neutralizing antibodies (Leu et al. (1998) *AIDS Res. and Human Retroviruses.* 14:151-155).

Further, a subset of the broadly reactive antibodies, found in most infected individuals, interferes with the binding of gp120 and CD4 (Kang, C.-Y., et al. (1991) *Proc. Natl. Acad. Sci. USA*. **88**:6171-6175; McDougal, J. S., et al. (1986) *J. Immunol.* **137**:2937-2944). Other antibodies are believed to bind to the chemokine receptor binding region after CD4 has bound to Env (Thali et al. (1993) *J. Virol.* 67:3978-3988). The fact that neutralizing

antibodies generated during the course of HIV infection do not provide permanent antiviral effect may in part be due to the generation of "neutralization escapes" virus mutants and to the general decline in the host immune system associated with pathogenesis. In contrast, the presence of pre-existing neutralizing antibodies upon initial HIV-1 exposure will likely have a protective effect.

5

10

15

20

25

30

It is widely thought that a successful vaccine should be able to induce a strong, broadly neutralizing antibody response against diverse HIV-1 strains (Montefiori and Evans (1999) AIDS Res. Hum. Ret. 15(8):689-698; Bolognesi, D.,P., et al. (1994) Ann. Int. Med. 8:603-611; Haynes, B., F., et al. (1996) Science; 271: 324-328.). Neutralizing antibodies, by attaching to the incoming virions, can reduce or even prevent their infectivity for target cells and prevent the cell-to-cell spread of virus in tissue culture (Hu et al. (1992) Science 255:456-459; Burton, D.,R. and Montefiori, D. (1997) AIDS 11(suppl. A): 587-598). However as described above, antibodies directed against gp120 do not generally exhibit broad antibody responses against different HIV strains.

Currently, the focus of vaccine development, from the perspective of humoral immunity, is on the neutralization of primary isolates that utilize the CCR5 chemokine coreceptor believed to be important in virus entry (Zhu, T., et al. (1993) *Science* 261:1179-1181; Fiore, J., et al. (1994) Virology; 204:297-303). These viruses are generally much more resistant to antibody neutralization than T-cell line adapted strains that use the CXCR4 coreceptor, although both can be neutralized *in vitro* by certain broadly and potent acting monoclonal antibodies, such as IgG1b12, 2G12 and 2F5 (Trkola, A., et al. (1995) *J. Virol.* 69:6609-6617; D'Sousa PM., et al (1997) *J. Infect. Dis.* 175:1062-1075). These monoclonal antibodies are directed to the CD4 binding site, a glycosylation site and to the gp41 fusion domain, respectively. The problem that remains, however, is that it is not known how to induce antibodies of the appropriate specificity by vaccination. Antibodies (Abs) elicited by gp120 glycoprotein from a given isolate are usually only able to neutralize closely related viruses generally from similar, usually from the same, HIV-1 subtype.

Despite the above approaches, there remains a need for Env antigens that can elicit an immunological response (e.g., neutralizing and/or protective antibodies) in a subject against multiple HIV strains and subtypes, for example when administered as a vaccine. The present invention solves these and other problems by providing modified Env polypeptides (e.g., gp120) to expose epitopes in or near the CD4 binding site.

Summary of the Invention

5

10

15

20

25

30

In accordance with the present invention, modified HIV Env polypeptides are provided. In particular, deletions and/or mutations are made in one or more of the $4-\beta$ antiparallel-bridging sheet in the HIV Env polypeptide. In this way, enough structure is left to allow correct folding of the polypeptide, for example of gp120, yet enough of the bridging sheet is removed to expose the CD4 groove, allowing an immune response to be generated against epitopes in or near the CD4 binding site of the Env polypeptide (e.g., gp120).

In one aspect, the invention includes a polynucleotide encoding a modified HIV Env polypeptide wherein the polypeptide has at least one modified (*e.g.*, deleted or replaced) amino acid residue deleted in the region corresponding to residues 421 to 436 relative to HXB-2, for example the constructs depicted in Figures 6-29 (SEQ ID NOs:3 to 26). In certain embodiments, the polynucleotide also has the region corresponding to residues 124-198 of the polypeptide HXB-2 (*e.g.*, V1/V2) deleted and at least one amino acid deleted or replaced in the regions corresponding to the residues 119 to 123 and 199 to 210, relative to HXB-2. In other embodiments, these polynucleotides encode Env polypeptides having at least one amino acid of the small loop of the bridging sheet (*e.g.*, amino acid residues 427 to 429 relative to HXB-2) deleted or replaced. The amino acid sequences of the modified polypeptides encoded by the polynucleotides of the present invention can be based on any HIV variant, for example SF162.

In another aspect, the invention includes immunogenic modified HIV Env polypeptides having at least one modified (*e.g.*, deleted or replaced) amino acid residue deleted in the region corresponding to residues 421 to 436 relative to HXB-2, for example a deletion or replacement of one amino acids in the small loop region (*e.g.*, amino acid residues 427 to 429 relative to HXB-2). These polypeptides may have modifications (*e.g.*, a deletion or a replacement) of at least one amino acid between about amino acid residue 420 and amino acid residue 436, relative to HXB-2 and, optionally, may have deletions or truncations of the V1 and/or V2 regions. The immunogenic, modified polypeptides of the present invention can be based on any HIV variant, for example SF162.

In another aspect, the invention includes a vaccine composition comprising any of the polynucleotides encoding modified Env polypeptides described above. Vaccine compositions comprising the modified Env polypeptides and, optionally, an adjuvant are also included in the invention.

In yet another aspect, the invention includes a method of inducing an immune response in subject comprising, administering one or more of the polynucleotides or constructs described above in an amount sufficient to induce an immune response in the subject. In certain embodiments, the method further comprises administering an adjuvant to the subject.

In another aspect, the invention includes a method of inducing an immune response in a subject comprising administering a composition comprising any of the modified Env polypeptides described above and an adjuvant. The composition is administered in an amount sufficient to induce an immune response in the subject.

In another aspect, the invention includes a method of inducing an immune response in a subject comprising

- (a) administering a first composition comprising any of the polynucleotides described above in a priming step and
- (b) administering a second composition comprising any of the modified Env polypeptides described above, as a booster, in an amount sufficient to induce an immune response in the subject. In certain embodiments, the first composition, the second composition or both the first and second compositions further comprise an adjuvant.

These and other embodiments of the subject invention will readily occur to those of skill in the art in light of the disclosure herein.

20

25

30

5

10

15

Brief Description of the Drawings

Figure 1 is a schematic depiction of the tertiary structure of the HIV- $1_{\rm HXB-2}$ Env gp120 polypeptide, as determined by crystallography studies.

Figures 2A-C depict alignment of the amino acid sequence of wild-type HIV-1_{HXB-2} Env gp160 polypeptide (SEQ ID NO:1) with amino acid sequence of HIV variants SF162 (shown as "162") (SEQ ID NO:2), SF2, CM236 and US4. Arrows indicate the regions that are deleted or replaced in the modified polypeptides. Black dots indicate conserved cysteine residues. The star indicates the position of the last amino acid in gp120.

Figures 3A-J depict alignment of nucleotide sequences of polynucleotides encoding modified Env polypeptides having V1/V2 deletions. The unmodified amino acid residues encoded by these sequences correspond to wildtype SF162 residues but are numbered relative to HXB-2.

Figures 4A-M depict alignment of nucleotide sequences of polynucleotides encoding modified Env polypeptides having deletions or replacements in the small loop. The unmodified amino acid residues encoded by these sequences correspond to wildtype SF162 residues but are numbered relative to HXB-2.

Figures 5A-N depict alignment of nucleotide sequences of polynucleotides encoding modified Env polypeptides having both V1/V2 deletions and, in addition, deletions or replacements in the small loop. The unmodified amino acid residues encoded by these sequences correspond to wildtype SF162 residues but are numbered relative to HXB-2.

5

10

20

25

Figure 6 depicts the nucleotide sequence of the construct designated Val120-Ala204 (SEQ ID NO:3).

Figure 7 depicts the nucleotide sequence of the construct designated Val120-Ile201 (SEQ ID NO:4).

Figure 8 depicts the nucleotide sequence of the construct designated Val120-Ile201B (SEO ID NO:5).

Figure 9 depicts the nucleotide sequence of the construct designated Lys121-Val200 (SEQ ID NO:6).

Figure 10 depicts the nucleotide sequence of the construct designated Leu122-Ser199 (SEQ ID NO:7).

Figure 11 depicts the nucleotide sequence of the construct designated Val120-Thr202 (SEQ ID NO:8).

Figure 12 depicts the nucleotide sequence of the construct designated Trp427-Gly431 (SEQ ID NO:9).

Figure 13 depicts the nucleotide sequence of the construct designated Arg426-Gly431 (SEQ ID NO:10).

Figure 14 depicts the nucleotide sequence of the construct designated Arg426-Gly431B (SEQ ID NO:11).

Figure 15 depicts the nucleotide sequence of the construct designated Arg426-Lys432 (SEQ ID NO:12).

Figure 16 depicts the nucleotide sequence of the construct designated Asn425-Lys432 (SEQ ID NO:13).

Figure 17 depicts the nucleotide sequence of the construct designated Ile424-Ala433 (SEQ ID NO:14).

Figure 18 depicts the nucleotide sequence of the construct designated Ile423-Met434 (SEQ ID NO:15).

Figure 19 depicts the nucleotide sequence of the construct designated Gln422-Tyr435 (SEQ ID NO:16).

Figure 20 depicts the nucleotide sequence of the construct designated Gln422-Tyr435B (SEQ ID NO:17).

Figure 21 depicts the nucleotide sequence of the construct designated Leu122-Ser199;Arg426-Gly431 (SEQ ID NO:18).

Figure 22 depicts the nucleotide sequence of the construct designated Leu122-Ser199;Arg426-Lys432 (SEQ ID NO:19).

Figure 23 depicts the nucleotide sequence of the construct designated Leu122-Ser199; Trp427-Gly431 (SEQ ID NO:20).

Figure 24 depicts the nucleotide sequence of the construct designated Lys121-Val200; Asn425-Lys432 (SEQ ID NO:21).

Figure 25 depicts the nucleotide sequence of the construct designated Val120-Ile201; Ile424-Ala433 (SEQ ID NO:22).

Figure 26 depicts the nucleotide sequence of the construct designated Val120-Ile201B; Ile424-Ala433 (SEQ ID NO:23).

Figure 27 depicts the nucleotide sequence of the construct designated Val120-Thr202; Ile424-Ala433 (SEQ ID NO:24).

Figure 28 depicts the nucleotide sequence of the construct designated Val127-Asn195 (SEQ ID NO:25).

Figure 29 depicts the nucleotide sequence of the construct designated Val127-Asn195; Arg426-Gly431 (SEQ ID NO:26).

25

30

20

10

Detailed Description of the Invention

The practice of the present invention will employ, unless otherwise indicated, conventional methods of protein chemistry, viral immunobiology, molecular biology and recombinant DNA techniques within the skill of the art. Such techniques are explained fully in the literature. See, e.g., T.E. Creighton, <u>Proteins: Structures and Molecular Properties</u> (W.H. Freeman and Company, 1993); Nelson L.M. and Jerome H.K. <u>HIV Protocols</u> in Methods in Molecular Medicine, vol. 17, 1999; Sambrook, et al., <u>Molecular Cloning: A</u>

<u>Laboratory Manual</u> (Cold Spring Harbor Laboratory, 1989); F.M. Ausubel et al. <u>Current Protocols in Molecular Biology</u>, Greene Publishing Associates & Wiley Interscience New York; and Lipkowitz and Boyd, <u>Reviews in Computational Chemistry</u>, volumes 1-present (Wiley-VCH, New York, New York, 1999).

It must be noted that, as used in this specification and the appended claims, the singular forms "a", "an" and "the" include plural referents unless the content clearly dictates otherwise. Thus, for example, reference to "a polypeptide" includes a mixture of two or more polypeptides, and the like.

10 **Definitions**

5

15

20

25

30

In describing the present invention, the following terms will be employed, and are intended to be defined as indicated below.

The terms "polypeptide," and "protein" are used interchangeably herein to denote any polymer of amino acid residues. The terms encompass peptides, oligopeptides, dimers, multimers, and the like. Such polypeptides can be derived from natural sources or can be synthesized or recombinantly produced. The terms also include postexpression modifications of the polypeptide, for example, glycosylation, acetylation, phosphorylation, etc.

A polypeptide as defined herein is generally made up of the 20 natural amino acids Ala (A), Arg (R), Asn (N), Asp (D), Cys (C), Gln (Q), Glu (E), Gly (G), His (H), Ile (I), Leu (L), Lys (K), Met (M), Phe (F), Pro (P), Ser (S), Thr (T), Trp (W), Tyr (Y) and Val (V) and may also include any of the several known amino acid analogs, both naturally occurring and synthesized analogs, such as but not limited to homoisoleucine, asaleucine, 2- (methylenecyclopropyl)glycine, S-methylcysteine, S-(prop-l-enyl)cysteine, homoserine, ornithine, norleucine, norvaline, homoarginine, 3-(3-carboxyphenyl)alanine, cyclohexylalanine, mimosine, pipecolic acid, 4-methylglutamic acid, canavanine, 2,3-diaminopropionic acid, and the like. Further examples of polypeptide agents which will find use in the present invention are set forth below.

By "geometry" or "tertiary structure" of a polypeptide or protein is meant the overall 3-D configuration of the protein. As described herein, the geometry can be determined, for example, by crystallography studies or by using various programs or algorithms which predict the geometry based on interactions between the amino acids making up the primary and secondary structures.

By "wild type" polypeptide, polypeptide agent or polypeptide drug, is meant a naturally occurring polypeptide sequence, and its corresponding secondary structure. An "isolated" or "purified" protein or polypeptide is a protein which is separate and discrete from a whole organism with which the protein is normally associated in nature. It is apparent that the term denotes proteins of various levels of purity. Typically, a composition containing a purified protein will be one in which at least about 35%, preferably at least about 40-50%, more preferably, at least about 75-85%, and most preferably at least about 90% or more, of the total protein in the composition will be the protein in question.

5

10

15

20

25

30

By "Env polypeptide" is meant a molecule derived from an envelope protein, preferably from HIV Env. The envelope protein of HIV-1 is a glycoprotein of about 160 kd (gp160). During virus infection of the host cell, gp160 is cleaved by host cell proteases to form gp120 and the integral membrane protein, gp41. The gp41 portion is anchored in (and spans) the membrane bilayer of virion, while the gp120 segment protrudes into the surrounding environment. As there is no covalent attachment between gp120 and gp41, free gp120 is released from the surface of virions and infected cells. Env polypeptides may also include gp140 polypeptides. Env polypeptides can exist as monomers, dimers or multimers.

By a "gp120 polypeptide" is meant a molecule derived from a gp120 region of the Env polypeptide. Preferably, the gp120 polypeptide is derived from HIV Env. The primary amino acid sequence of gp120 is approximately 511 amino acids, with a polypeptide core of about 60,000 daltons. The polypeptide is extensively modified by N-linked glycosylation to increase the apparent molecular weight of the molecule to 120,000 daltons. The amino acid sequence of gp120 contains five relatively conserved domains interspersed with five hypervariable domains. The positions of the 18 cysteine residues in the gp120 primary sequence of the HIV-1_{HXB-2} (hereinafter "HXB-2") strain, and the positions of 13 of the approximately 24 N-linked glycosylation sites in the gp120 sequence are common to most, if not all, gp120 sequences. The hypervariable domains contain extensive amino acid substitutions, insertions and deletions. Despite this variation, most, if not all, gp120 sequences preserve the virus's ability to bind to the viral receptor CD4. A "gp120 polypeptide" includes both single subunits or multimers.

Env polypeptides (e.g., gp120, gp140 and gp160) include a "bridging sheet" comprised of 4 anti-parallel β -strands (β -2, β -3, β -20 and β -21) that form a β -sheet. Extruding from one pair of the β -strands (β -2 and β -3) are two loops, V1 and V2. The β -2

sheet occurs at approximately amino acid residue 119 (Cys) to amino acid residue 123 (Thr) while β -3 occurs at approximately amino acid residue 199 (Ser) to amino acid residue 201 (Ile), relative to HXB-2. The "V1/V2 region" occurs at approximately amino acid positions 126 (Cys) to residue 196 (Cys), relative to HXB-2. (see, *e.g.*, Wyatt et al. (1995) *J. Virol.* 69:5723-5733; Stamatatos et al. (1998) *J. Virol.* 72:7840-7845). Extruding from the second pair of β -strands (β -20 and β -21) is a "small-loop" structure, also referred to herein as "the bridging sheet small loop." In HXB-2, β -20 extends from about amino acid residue 422 (Gln) to amino acid residue 426 (Met) while β -21 extends from about amino acid residue 430 (Val) to amino acid residue 435 (Tyr). In variant SF162, the Met-426 is an Arg (R) residue. The "small loop" extends from about amino acid residue 427 (Trp) through 429 (Lys), relative to HXB-2. A representative diagram of gp120 showing the bridging sheet, the small loop, and V1/V2 is shown in Figure 1. In addition, alignment of the amino acid sequences of Env polypeptide gp160 of selected variants is shown, relative to HXB-2, in Figures 2A-C.

5

10

15

20

25

30

Furthermore, an "Env polypeptide" or "gp120 polypeptide" as defined herein is not limited to a polypeptide having the exact sequence described herein. Indeed, the HIV genome is in a state of constant flux and contains several variable domains which exhibit relatively high degrees of variability between isolates. It is readily apparent that the terms encompass Env (e.g., gp120) polypeptides from any of the identified HIV isolates, as well as newly identified isolates, and subtypes of these isolates. Descriptions of structural features are given herein with reference to HXB-2. One of ordinary skill in the art in view of the teachings of the present disclosure and the art can determine corresponding regions in other HIV variants (e.g., isolates HIV_{IIIb}, HIV_{SE2}, HIV-1_{SE162}, HIV-1_{SE170}, HIV_{LAV}, HIV_{LAI}, HIV_{MN}, HIV-1_{CM235}, HIV-1_{US4}, other HIV-1 strains from diverse subtypes(e.g., subtypes, A through G, and O), HIV-2 strains and diverse subtypes (e.g., HIV- 2_{UC1} and HIV- 2_{UC2}), and simian immunodeficiency virus (SIV). (See, e.g., Virology, 3rd Edition (W.K. Joklik ed. 1988); Fundamental Virology, 2nd Edition (B.N. Fields and D.M. Knipe, eds. 1991); Virology, 3rd Edition (Fields, BN, DM Knipe, PM Howley, Editors, 1996, Lippincott-Raven, Philadelphia, PA; for a description of these and other related viruses), using for example, sequence comparison programs (e.g., BLAST and others described herein) or identification and alignment of structural features (e.g., a program such as the "ALB" program described herein that can identify \(\theta\)-sheet regions). The actual amino acid sequences of the modified Env polypeptides can be based on any HIV variant.

Additionally, the term "Env polypeptide" (e.g., "gp120 polypeptide") encompasses proteins which include additional modifications to the native sequence, such as additional internal deletions, additions and substitutions. These modifications may be deliberate, as through site-directed mutagenesis, or may be accidental, such as through naturally occurring mutational events. Thus, for example, if the Env polypeptide is to be used in vaccine compositions, the modifications must be such that immunological activity (i.e., the ability to elicit an antibody response to the polypeptide) is not lost. Similarly, if the polypeptides are to be used for diagnostic purposes, such capability must be retained.

5

10

15

20

25

30

Thus, a "modified Env polypeptide" is an Env polypeptide (*e.g.*, gp120 as defined above), which has been manipulated to delete or replace all or a part of the bridging sheet portion and, optionally, the variable regions V1 and V2. Generally, modified Env (*e.g.*, gp120) polypeptides have enough of the bridging sheet removed to expose the CD4 binding site, but leave enough of the structure to allow correct folding (*e.g.*, correct geometry). Thus, modifications to the β -20 and β -21 regions (between about amino acid residues 420 and 435 relative to HXB-2) are preferred. Additionally, modifications to the β -2 and β -3 regions (between about amino acid residues 119 (Cys) and 201 (Ile)) and modifications (*e.g.*, truncations) to the V1 and V2 loop regions may also be made. Although not all possible β -sheet and V1/V2 modifications have been exemplified herein, it is to be understood that other disrupting modifications are also encompassed by the present invention.

Normally, such a modified polypeptide is capable of secretion into growth medium in which an organism expressing the protein is cultured. However, for purposes of the present invention, such polypeptides may also be recovered intracellularly. Secretion into growth media is readily determined using a number of detection techniques, including, e.g., polyacrylamide gel electrophoresis and the like, and immunological techniques such as Western blotting and immunoprecipitation assays as described in, e.g., International Publication No. WO 96/04301, published February 15, 1996.

A gp120 or other Env polypeptide is produced "intracellularly" when it is found within the cell, either associated with components of the cell, such as in association with the endoplasmic reticulum (ER) or the Golgi Apparatus, or when it is present in the soluble cellular fraction. The gp120 and other Env polypeptides of the present invention may also be secreted into growth medium so long as sufficient amounts of the polypeptides remain

present within the cell such that they can be purified from cell lysates using techniques described herein.

5

10

15

20

25

30

An "immunogenic" gp120 or other Env protein is a molecule that includes at least one epitope such that the molecule is capable of either eliciting an immunological reaction in an individual to which the protein is administered or, in the diagnostic context, is capable of reacting with antibodies directed against the HIV in question.

By "epitope" is meant a site on an antigen to which specific B cells and/or T cells respond, rendering the molecule including such an epitope capable of eliciting an immunological reaction or capable of reacting with HIV antibodies present in a biological sample. The term is also used interchangeably with "antigenic determinant" or "antigenic determinant site." An epitope can comprise 3 or more amino acids in a spatial conformation unique to the epitope. Generally, an epitope consists of at least 5 such amino acids and, more usually, consists of at least 8-10 such amino acids. Methods of determining spatial conformation of amino acids are known in the art and include, for example, x-ray crystallography and 2-dimensional nuclear magnetic resonance. Furthermore, the identification of epitopes in a given protein is readily accomplished using techniques well known in the art, such as by the use of hydrophobicity studies and by site-directed serology. See, also, Geysen et al., Proc. Natl. Acad. Sci. USA (1984) 81:3998-4002 (general method of rapidly synthesizing peptides to determine the location of immunogenic epitopes in a given antigen); U.S. Patent No. 4,708,871 (procedures for identifying and chemically synthesizing epitopes of antigens); and Geysen et al., Molecular Immunology (1986) 23:709-715 (technique for identifying peptides with high affinity for a given antibody). Antibodies that recognize the same epitope can be identified in a simple immunoassay showing the ability of one antibody to block the binding of another antibody to a target antigen.

An "immunological response" or "immune response" as used herein is the development in the subject of a humoral and/or a cellular immune response to the Env (e.g., gp120) polypeptide when the polypeptide is present in a vaccine composition. These antibodies may also neutralize infectivity, and/or mediate antibody-complement or antibody dependent cell cytotoxicity to provide protection to an immunized host. Immunological reactivity may be determined in standard immunoassays, such as a competition assays, well known in the art.

Techniques for determining amino acid sequence "similarity" are well known in the art. In general, "similarity" means the exact amino acid to amino acid comparison of two or more polypeptides at the appropriate place, where amino acids are identical or possess similar chemical and/or physical properties such as charge or hydrophobicity. A so-termed "percent similarity" then can be determined between the compared polypeptide sequences.

Techniques for determining nucleic acid and amino acid sequence identity also are well known in the art and include determining the nucleotide sequence of the mRNA for that gene (usually via a cDNA intermediate) and determining the amino acid sequence encoded thereby, and comparing this to a second amino acid sequence. In general, "identity" refers to an exact nucleotide to nucleotide or amino acid to amino acid correspondence of two polynucleotides or polypeptide sequences, respectively.

5

10

15

20

25

30

Two or more polynucleotide sequences can be compared by determining their "percent identity." Two or more amino acid sequences likewise can be compared by determining their "percent identity." The percent identity of two sequences, whether nucleic acid or peptide sequences, is generally described as the number of exact matches between two aligned sequences divided by the length of the shorter sequence and multiplied by 100. An approximate alignment for nucleic acid sequences is provided by the local homology algorithm of Smith and Waterman, Advances in Applied Mathematics 2:482-489 (1981). This algorithm can be extended to use with peptide sequences using the scoring matrix developed by Dayhoff, Atlas of Protein Sequences and Structure, M.O. Dayhoff ed., 5 suppl. 3:353-358, National Biomedical Research Foundation, Washington, D.C., USA, and normalized by Gribskov, Nucl. Acids Res. 14(6):6745-6763 (1986). An implementation of this algorithm for nucleic acid and peptide sequences is provided by the Genetics Computer Group (Madison, WI) in their BestFit utility application. The default parameters for this method are described in the Wisconsin Sequence Analysis Package Program Manual, Version 8 (1995) (available from Genetics Computer Group, Madison, WI). Other equally suitable programs for calculating the percent identity or similarity between sequences are generally known in the art.

For example, percent identity of a particular nucleotide sequence to a reference sequence can be determined using the homology algorithm of Smith and Waterman with a default scoring table and a gap penalty of six nucleotide positions. Another method of establishing percent identity in the context of the present invention is to use the MPSRCH

package of programs copyrighted by the University of Edinburgh, developed by John F. Collins and Shane S. Sturrok, and distributed by IntelliGenetics, Inc. (Mountain View, CA). From this suite of packages, the Smith-Waterman algorithm can be employed where default parameters are used for the scoring table (for example, gap open penalty of 12, gap extension penalty of one, and a gap of six). From the data generated, the "Match" value reflects "sequence identity." Other suitable programs for calculating the percent identity or similarity between sequences are generally known in the art, such as the alignment program BLAST, which can also be used with default parameters. For example, BLASTN and BLASTP can be used with the following default parameters: genetic code = standard; filter = none; strand = both; cutoff = 60; expect = 10; Matrix = BLOSUM62; Descriptions = 50 sequences; sort by = HIGH SCORE; Databases = non-redundant, GenBank + EMBL + DDBJ + PDB + GenBank CDS translations + Swiss protein + Spupdate + PIR. Details of these programs can be found at the following internet address: http://www.ncbi.nlm.gov/cgi-bin/BLAST.

One of skill in the art can readily determine the proper search parameters to use for a given sequence in the above programs. For example, the search parameters may vary based on the size of the sequence in question. Thus, for example, a representative embodiment of the present invention would include an isolated polynucleotide having X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least about 50% identity to Y contiguous nucleotides derived from any of the sequences described herein, (ii) X equals Y, and (iii) X is greater than or equal to 6 nucleotides and up to 5000 nucleotides, preferably greater than or equal to 8 nucleotides and up to 5000 nucleotides, more preferably 10-12 nucleotides and up to 5000 nucleotides, and even more preferably 15-20 nucleotides, up to the number of nucleotides present in the full-length sequences described herein (e.g., see the Sequence Listing and claims), including all integer values falling within the above-described ranges.

The synthetic expression cassettes (and purified polynucleotides) of the present invention include related polynucleotide sequences having about 80% to 100%, greater than 80-85%, preferably greater than 90-92%, more preferably greater than 95%, and most preferably greater than 98% sequence (including all integer values falling within these described ranges) identity to the synthetic expression cassette sequences disclosed herein (for example, to the claimed sequences or other sequences of the present invention) when the sequences of the present invention are used as the query sequence.

Computer programs are also available to determine the likelihood of certain polypeptides to form structures such as β-sheets. One such program, described herein, is the "ALB" program for protein and polypeptide secondary structure calculation and predication. In addition, secondary protein structure can be predicted from the primary amino acid sequence, for example using protein crystal structure and aligning the protein sequence related to the crystal structure (*e.g.*, using Molecular Operating Environment (MOE) programs available from the Chemical Computing Group Inc., Montreal, P.Q., Canada). Other methods of predicting secondary structures are described, for example, in Garnier et al. (1996) *Methods Enzymol.* 266:540-553; Geourjon et al. (1995) *Comput. Applic. Biosci.* 11:681-684; Levin (1997) *Protein Eng.* 10:771-776; and Rost et al. (1993) *J. Molec. Biol.* 232:584-599.

Homology can also be determined by hybridization of polynucleotides under conditions which form stable duplexes between homologous regions, followed by digestion with single-stranded-specific nuclease(s), and size determination of the digested fragments. Two DNA, or two polypeptide sequences are "substantially homologous" to each other when the sequences exhibit at least about 80%-85%, preferably at least about 90%, and most preferably at least about 95%-98% sequence identity over a defined length of the molecules, as determined using the methods above. As used herein, substantially homologous also refers to sequences showing complete identity to the specified DNA or polypeptide sequence. DNA sequences that are substantially homologous can be identified in a Southern hybridization experiment under, for example, stringent conditions, as defined for that particular system. Defining appropriate hybridization conditions is within the skill of the art. See, e.g., Sambrook et al., supra; DNA Cloning, supra; Nucleic Acid Hybridization, supra.

A "coding sequence" or a sequence which "encodes" a selected protein, is a nucleic acid sequence which is transcribed (in the case of DNA) and translated (in the case of mRNA) into a polypeptide *in vitro* or *in vivo* when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a start codon at the 5' (amino) terminus and a translation stop codon at the 3' (carboxy) terminus. A coding sequence can include, but is not limited to cDNA from viral nucleotide sequences as well as synthetic and semisynthetic DNA sequences and sequences including base analogs. A transcription termination sequence may be located 3' to the coding sequence.

"Control elements" refers collectively to promoter sequences, ribosome binding sites, polyadenylation signals, transcription termination sequences, upstream regulatory domains, enhancers, and the like, which collectively provide for the transcription and translation of a coding sequence in a host cell. Not all of these control elements need always be present so long as the desired gene is capable of being transcribed and translated.

5

10

15

20

25

30

A control element "directs the transcription" of a coding sequence in a cell when RNA polymerase will bind the promoter sequence and transcribe the coding sequence into mRNA, which is then translated into the polypeptide encoded by the coding sequence.

"Operably linked" refers to an arrangement of elements wherein the components so described are configured so as to perform their usual function. Thus, control elements operably linked to a coding sequence are capable of effecting the expression of the coding sequence when RNA polymerase is present. The control elements need not be contiguous with the coding sequence, so long as they function to direct the expression thereof. Thus, for example, intervening untranslated yet transcribed sequences can be present between, c.g., a promoter sequence and the coding sequence and the promoter sequence can still be considered "operably linked" to the coding sequence.

"Recombinant" as used herein to describe a nucleic acid molecule means a polynucleotide of genomic, cDNA, semisynthetic, or synthetic origin which, by virtue of its origin or manipulation: (1) is not associated with all or a portion of the polynucleotide with which it is associated in nature; and/or (2) is linked to a polynucleotide other than that to which it is linked in nature. The term "recombinant" as used with respect to a protein or polypeptide means a polypeptide produced by expression of a recombinant polynucleotide. "Recombinant host cells," "host cells," "cells," "cell lines," "cell cultures," and other such terms denoting procaryotic microorganisms or eucaryotic cell lines cultured as unicellular entities, are used interchangeably, and refer to cells which can be, or have been, used as recipients for recombinant vectors or other transfer DNA, and include the progeny of the original cell which has been transfected. It is understood that the progeny of a single parental cell may not necessarily be completely identical in morphology or in genomic or total DNA complement to the original parent, due to accidental or deliberate mutation. Progeny of the parental cell which are sufficiently similar to the parent to be characterized by the relevant property, such as the presence of a nucleotide sequence encoding a desired peptide, are included in the progeny intended by this definition, and are covered by the above terms.

By "vertebrate subject" is meant any member of the subphylum chordata, including, without limitation, humans and other primates, including non-human primates such as chimpanzees and other apes and monkey species; farm animals such as cattle, sheep, pigs, goats and horses; domestic mammals such as dogs and cats; laboratory animals including rodents such as mice, rats and guinea pigs; birds, including domestic, wild and game birds such as chickens, turkeys and other gallinaceous birds, ducks, geese, and the like. The term does not denote a particular age. Thus, both adult and newborn individuals are intended to be covered.

As used herein, a "biological sample" refers to a sample of tissue or fluid isolated from an individual, including but not limited to, for example, blood, plasma, serum, fecal matter, urine, bone marrow, bile, spinal fluid, lymph fluid, samples of the skin, external secretions of the skin, respiratory, intestinal, and genitourinary tracts, samples derived from the gastric epithelium and gastric mucosa, tears, saliva, milk, blood cells, organs, biopsies and also samples of *in vitro* cell culture constituents including but not limited to conditioned media resulting from the growth of cells and tissues in culture medium, e.g., recombinant cells, and cell components.

The terms "label" and "detectable label" refer to a molecule capable of detection, including, but not limited to, radioactive isotopes, fluorescers, chemiluminescers, enzymes, enzyme substrates, enzyme cofactors, enzyme inhibitors, chromophores, dyes, metal ions, metal sols, ligands (e.g., biotin or haptens) and the like. The term "fluorescer" refers to a substance or a portion thereof which is capable of exhibiting fluorescence in the detectable range. Particular examples of labels which may be used with the invention include, but are not limited to fluorescein, rhodamine, dansyl, umbelliferone, Texas red, luminol, acradimum esters, NADPH, α - β -galactosidase, horseradish peroxidase, glucose oxidase, alkaline phosphatase and urease.

Overview

5

10

15

20

25

30

The present invention concerns modified Env polypeptide molecules (e.g., glycoprotein ("gp") 120). Without being bound by a particular theory, it appears that it has been difficult to generate immunological responses against Env because the CD4 binding site is buried between the outer domain, the inner domain and the V1/V2 domains. Thus, although deletion of the V1/V2 domain may render the virus more susceptible to

neutralization by monoclonal antibody directed to the CD4 site, the bridging sheet covering most of the CD4 binding domain may prevent an antibody response. Thus, the present invention provides Env polypeptides that maintain their general overall structure yet expose the CD4 binding domain. This allows the generation of an immune response (*e.g.*, an antibody response) to epitopes in or near the CD4 binding site.

Various forms of the different embodiments of the invention, described herein, may be combined.

β-Sheet Conformations

5

10

15

20

25

30

In the present invention, location of the β -sheet structures were identified relative to 3-D (crystal) structure of an HXB-2 crystallized Env protein (see, Example 1A). Based on this structure, constructs encoding polypeptides having replacements and or excisions which maintain overall geometry while exposing the CD4 binding site were designed. In particular, the crystal structure of HXB-2 was downloaded from the Brookhaven Database. Using the default parameters of the Loop Search feature of the Biopolymer module of the Sybyl molecular modeling package, homology and fit of amino acids which could replace the native loops between β -strands yet maintain overall tertiary structure were determined. Constructs encoding the modified Env polypeptides were then designed (Example 1.B.).

Thus, the modified Env polypeptides typically have enough of the bridging sheet removed to expose the CD4 groove, but have enough of the structure to allow correct folding of the Env glycoprotein. Exemplary constructs are described below.

Polypeptide Production

The polypeptides of the present invention can be produced in any number of ways which are well known in the art.

In one embodiment, the polypeptides are generated using recombinant techniques, well known in the art. In this regard, oligonucleotide probes can be devised based on the known sequences of the Env (e.g., gp120) polypeptide genome and used to probe genomic or cDNA libraries for Env genes. The gene can then be further isolated using standard techniques and, e.g., restriction enzymes employed to truncate the gene at desired portions of the full-length sequence. Similarly, the Env gene(s) can be isolated directly from cells and tissues containing the same, using known techniques, such as phenol extraction and the

sequence further manipulated to produce the desired truncations. *See, e.g.,* Sambrook et al., *supra*, for a description of techniques used to obtain and isolate DNA.

The genes encoding the modified (e.g., truncated and/or substituted) polypeptides can be produced synthetically, based on the known sequences. The nucleotide sequence can be designed with the appropriate codons for the particular amino acid sequence desired. The complete sequence is generally assembled from overlapping oligonucleotides prepared by standard methods and assembled into a complete coding sequence. See, e.g., Edge (1981) Nature 292:756; Nambair et al. (1984) Science 223:1299; Jay et al. (1984) J. Biol. Chem. 259:6311; Stemmer et al. (1995) Gene 164:49-53.

5

10

15

20

25

30

Recombinant techniques are readily used to clone a gene encoding an Env polypeptide gene which can then be mutagenized *in vitro* by the replacement of the appropriate base pair(s) to result in the codon for the desired amino acid. Such a change can include as little as one base pair, effecting a change in a single amino acid, or can encompass several base pair changes. Alternatively, the mutations can be effected using a mismatched primer which hybridizes to the parent nucleotide sequence (generally cDNA corresponding to the RNA sequence), at a temperature below the melting temperature of the mismatched duplex. The primer can be made specific by keeping primer length and base composition within relatively narrow limits and by keeping the mutant base centrally located. See, *e.g.*, Innis et al, (1990) PCR Applications: Protocols for Functional Genomics; Zoller and Smith, *Methods Enzymol.* (1983) 100:468. Primer extension is effected using DNA polymerase, the product cloned and clones containing the mutated DNA, derived by segregation of the primer extended strand, selected. Selection can be accomplished using the mutant primer as a hybridization probe. The technique is also applicable for generating multiple point mutations. See, e.g., Dalbie-McFarland et al. *Proc. Natl. Acad. Sci USA* (1982) 79:6409.

Once coding sequences for the desired proteins have been isolated or synthesized, they can be cloned into any suitable vector or replicon for expression. As will be apparent from the teachings herein, a wide variety of vectors encoding modified polypeptides can be generated by creating expression constructs which operably link, in various combinations, polynucleotides encoding Env polypeptides having deletions or mutation therein. Thus, polynucleotides encoding a particular deleted V1/V2 region can be operably linked with polynucleotides encoding polypeptides having deletions or replacements in the small loop

region and the construct introduced into a host cell for polypeptide expression. Non-limiting examples of such combinations are discussed in the Examples.

Numerous cloning vectors are known to those of skill in the art, and the selection of an appropriate cloning vector is a matter of choice. Examples of recombinant DNA vectors for cloning and host cells which they can transform include the bacteriophage λ (*E. coli*), pBR322 (*E. coli*), pACYC177 (*E. coli*), pKT230 (gram-negative bacteria), pGV1106 (gram-negative bacteria), pLAFR1 (gram-negative bacteria), pME290 (non-*E. coli* gram-negative bacteria), pHV14 (*E. coli* and *Bacillus subtilis*), pBD9 (*Bacillus*), pIJ61 (*Streptomyces*), pUC6 (*Streptomyces*), YIp5 (*Saccharomyces*), YCp19 (*Saccharomyces*) and bovine papilloma virus (mammalian cells). *See*, *generally*, *DNA Cloning*: Vols. I & II, *supra*; Sambrook *et al.*, *supra*; B. Perbal, *supra*.

5

10

15

20

25

30

Insect cell expression systems, such as baculovirus systems, can also be used and are known to those of skill in the art and described in, *e.g.*, Summers and Smith, *Texas Agricultural Experiment Station Bulletin No. 1555* (1987). Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *inter alia*, Invitrogen, San Diego CA ("MaxBac" kit).

Plant expression systems can also be used to produce the modified Env proteins. Generally, such systems use virus-based vectors to transfect plant cells with heterologous genes. For a description of such systems see, e.g., Porta et al., *Mol. Biotech.* (1996) <u>5</u>:209-221; and Hackland et al., *Arch. Virol.* (1994) <u>139</u>:1-22.

Viral systems, such as a vaccinia based infection/transfection system, as described in Tomei et al., *J. Virol.* (1993) 67:4017-4026 and Selby et al., *J. Gen. Virol.* (1993) 74:1103-1113, will also find use with the present invention. In this system, cells are first transfected *in vitro* with a vaccinia virus recombinant that encodes the bacteriophage T7 RNA polymerase. This polymerase displays exquisite specificity in that it only transcribes templates bearing T7 promoters. Following infection, cells are transfected with the DNA of interest, driven by a T7 promoter. The polymerase expressed in the cytoplasm from the vaccinia virus recombinant transcribes the transfected DNA into RNA which is then translated into protein by the host translational machinery. The method provides for high level, transient, cytoplasmic production of large quantities of RNA and its translation product(s).

The gene can be placed under the control of a promoter, ribosome binding site (for bacterial expression) and, optionally, an operator (collectively referred to herein as "control" elements), so that the DNA sequence encoding the desired Env polypeptide is transcribed into RNA in the host cell transformed by a vector containing this expression construction. The coding sequence may or may not contain a signal peptide or leader sequence. With the present invention, both the naturally occurring signal peptides or heterologous sequences can be used. Leader sequences can be removed by the host in post-translational processing. *See*, *e.g.*, U.S. Patent Nos. 4,431,739; 4,425,437; 4,338,397. Such sequences include, but are not limited to, the TPA leader, as well as the honey bee mellitin signal sequence.

5

10

15

20

25

30

Other regulatory sequences may also be desirable which allow for regulation of expression of the protein sequences relative to the growth of the host cell. Such regulatory sequences are known to those of skill in the art, and examples include those which cause the expression of a gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound. Other types of regulatory elements may also be present in the vector, for example, enhancer sequences.

The control sequences and other regulatory sequences may be ligated to the coding sequence prior to insertion into a vector. Alternatively, the coding sequence can be cloned directly into an expression vector which already contains the control sequences and an appropriate restriction site.

In some cases it may be necessary to modify the coding sequence so that it may be attached to the control sequences with the appropriate orientation; *i.e.*, to maintain the proper reading frame. Mutants or analogs may be prepared by the deletion of a portion of the sequence encoding the protein, by insertion of a sequence, and/or by substitution of one or more nucleotides within the sequence. Techniques for modifying nucleotide sequences, such as site-directed mutagenesis, are well known to those skilled in the art. *See*, *e.g.*, Sambrook *et al.*, *supra*; *DNA Cloning*, Vols. I and II, *supra*; *Nucleic Acid Hybridization*, *supra*.

The expression vector is then used to transform an appropriate host cell. A number of mammalian cell lines are known in the art and include immortalized cell lines available from the American Type Culture Collection (ATCC), such as, but not limited to, Chinese hamster ovary (CHO) cells, HeLa cells, baby hamster kidney (BHK) cells, monkey kidney cells (COS), human hepatocellular carcinoma cells (e.g., Hep G2), Vero293 cells, as well as others. Similarly, bacterial hosts such as *E. coli*, *Bacillus subtilis*, and *Streptococcus spp.*, will find

use with the present expression constructs. Yeast hosts useful in the present invention include inter alia, Saccharomyces cerevisiae, Candida albicans, Candida maltosa, Hansenula polymorpha, Kluyveromyces fragilis, Kluyveromyces lactis, Pichia guillerimondii, Pichia pastoris, Schizosaccharomyces pombe and Yarrowia lipolytica. Insect cells for use with baculovirus expression vectors include, inter alia, Aedes aegypti, Autographa californica, Bombyx mori, Drosophila melanogaster, Spodoptera frugiperda, and Trichoplusia ni.

5

10

15

20

25

30

Depending on the expression system and host selected, the proteins of the present invention are produced by growing host cells transformed by an expression vector described above under conditions whereby the protein of interest is expressed. The selection of the appropriate growth conditions is within the skill of the art.

In one embodiment, the transformed cells secrete the polypeptide product into the surrounding media. Certain regulatory sequences can be included in the vector to enhance secretion of the protein product, for example using a tissue plasminogen activator (TPA) leader sequence, a γ-interferon signal sequence or other signal peptide sequences from known secretory proteins. The secreted polypeptide product can then be isolated by various techniques described herein, for example, using standard purification techniques such as but not limited to, hydroxyapatite resins, column chromatography, ion-exchange chromatography, size-exclusion chromatography, electrophoresis, HPLC, immunoadsorbent techniques, affinity chromatography, immunoprecipitation, and the like..

Alternatively, the transformed cells are disrupted, using chemical, physical or mechanical means, which lyse the cells yet keep the Env polypeptides substantially intact. Intracellular proteins can also be obtained by removing components from the cell wall or membrane, e.g., by the use of detergents or organic solvents, such that leakage of the Env polypeptides occurs. Such methods are known to those of skill in the art and are described in, e.g., *Protein Purification Applications: A Practical Approach*, (E.L.V. Harris and S. Angal, Eds., 1990)

For example, methods of disrupting cells for use with the present invention include but are not limited to: sonication or ultrasonication; agitation; liquid or solid extrusion; heat treatment; freeze-thaw; desiccation; explosive decompression; osmotic shock; treatment with lytic enzymes including proteases such as trypsin, neuraminidase and lysozyme; alkali treatment; and the use of detergents and solvents such as bile salts, sodium dodecylsulphate,

Triton, NP40 and CHAPS. The particular technique used to disrupt the cells is largely a matter of choice and will depend on the cell type in which the polypeptide is expressed, culture conditions and any pre-treatment used.

Following disruption of the cells, cellular debris is removed, generally by centrifugation, and the intracellularly produced Env polypeptides are further purified, using standard purification techniques such as but not limited to, column chromatography, ion-exchange chromatography, size-exclusion chromatography, electrophoresis, HPLC, immunoadsorbent techniques, affinity chromatography, immunoprecipitation, and the like.

5

10

15

20

25

30

For example, one method for obtaining the intracellular Env polypeptides of the present invention involves affinity purification, such as by immunoaffinity chromatography using anti-Env specific antibodies, or by lectin affinity chromatography. Particularly preferred lectin resins are those that recognize mannose moieties such as but not limited to resins derived from *Galanthus nivalis* agglutinin (GNA), *Lens culinaris* agglutinin (LCA or lentil lectin), *Pisum sativum* agglutinin (PSA or pea lectin), *Narcissus pseudonarcissus* agglutinin (NPA) and *Allium ursinum* agglutinin (AUA). The choice of a suitable affinity resin is within the skill in the art. After affinity purification, the Env polypeptides can be further purified using conventional techniques well known in the art, such as by any of the techniques described above.

It may be desirable to produce Env (*e.g.*, gp120) complexes, either with itself or other proteins. Such complexes are readily produced by e.g., co-transfecting host cells with constructs encoding for the Env (*e.g.*, gp120) and/or other polypeptides of the desired complex. Co-transfection can be accomplished either in *trans* or *cis*, i.e., by using separate vectors or by using a single vector which bears both of the Env and other gene. If done using a single vector, both genes can be driven by a single set of control elements or, alternatively, the genes can be present on the vector in individual expression cassettes, driven by individual control elements. Following expression, the proteins will spontaneously associate. Alternatively, the complexes can be formed by mixing the individual proteins together which have been produced separately, either in purified or semi-purified form, or even by mixing culture media in which host cells expressing the proteins, have been cultured. See, International Publication No. WO 96/04301, published February 15, 1996, for a description of such complexes.

5

10

15

20

30

Relatively small polypeptides, i.e., up to about 50 amino acids in length, can be conveniently synthesized chemically, for example by any of several techniques that are known to those skilled in the peptide art. In general, these methods employ the sequential addition of one or more amino acids to a growing peptide chain. Normally, either the amino or carboxyl group of the first amino acid is protected by a suitable protecting group. The protected or derivatized amino acid can then be either attached to an inert solid support or utilized in solution by adding the next amino acid in the sequence having the complementary (amino or carboxyl) group suitably protected, under conditions that allow for the formation of an amide linkage. The protecting group is then removed from the newly added amino acid residue and the next amino acid (suitably protected) is then added, and so forth. After the desired amino acids have been linked in the proper sequence, any remaining protecting groups (and any solid support, if solid phase synthesis techniques are used) are removed sequentially or concurrently, to render the final polypeptide. By simple modification of this general procedure, it is possible to add more than one amino acid at a time to a growing chain, for example, by coupling (under conditions which do not racemize chiral centers) a protected tripeptide with a properly protected dipeptide to form, after deprotection, a pentapeptide. See, e.g., J. M. Stewart and J. D. Young, Solid Phase Peptide Synthesis (Pierce Chemical Co., Rockford, IL 1984) and G. Barany and R. B. Merrifield, The Peptides: Analysis, Synthesis, Biology, editors E. Gross and J. Meienhofer, Vol. 2, (Academic Press, New York, 1980), pp. 3-254, for solid phase peptide synthesis techniques; and M. Bodansky, Principles of Peptide Synthesis, (Springer-Verlag, Berlin 1984) and E. Gross and J. Meienhofer, Eds., The Peptides: Analysis, Synthesis, Biology, Vol. 1, for classical solution synthesis.

Typical protecting groups include t-butyloxycarbonyl (Boc), 9-

fluorenylmethoxycarbonyl (Fmoc) benzyloxycarbonyl (Cbz); p-toluenesulfonyl (Tx); 2,4-dinitrophenyl; benzyl (Bzl); biphenylisopropyloxycarboxy-carbonyl, t-amyloxycarbonyl, isobornyloxycarbonyl, o-bromobenzyloxycarbonyl, cyclohexyl, isopropyl, acetyl, o-nitrophenylsulfonyl and the like.

Typical solid supports are cross-linked polymeric supports. These can include divinylbenzene cross-linked-styrene-based polymers, for example, divinylbenzene-hydroxymethylstyrene copolymers, divinylbenzene-chloromethylstyrene copolymers and divinylbenzene-benzhydrylaminopolystyrene copolymers.

The polypeptide analogs of the present invention can also be chemically prepared by other methods such as by the method of simultaneous multiple peptide synthesis. See, e.g., Houghten *Proc. Natl. Acad. Sci. USA* (1985) 82:5131-5135; U.S. Patent No. 4,631,211.

5 Diagnostic and Vaccine Applications

10

15

20

25

30

The intracellularly produced Env polypeptides of the present invention, complexes thereof, or the polynucleotides coding therefor, can be used for a number of diagnostic and therapeutic purposes. For example, the proteins and polynucleotides or antibodies generated against the same, can be used in a variety of assays, to determine the presence of reactive antibodies/and or Env proteins in a biological sample to aid in the diagnosis of HIV infection or disease status or as measure of response to immunization.

The presence of antibodies reactive with the Env (e.g., gp120) polypeptides and, conversely, antigens reactive with antibodies generated thereto, can be detected using standard electrophoretic and immunodiagnostic techniques, including immunoassays such as competition, direct reaction, or sandwich type assays. Such assays include, but are not limited to, western blots; agglutination tests; enzyme-labeled and mediated immunoassays, such as ELISAs; biotin/avidin type assays; radioimmunoassays; immunoelectrophoresis; immunoprecipitation, etc. The reactions generally include revealing labels such as fluorescent, chemiluminescent, radioactive, or enzymatic labels or dye molecules, or other methods for detecting the formation of a complex between the antigen and the antibody or antibodies reacted therewith.

Solid supports can be used in the assays such as nitrocellulose, in membrane or microtiter well form; polyvinylchloride, in sheets or microtiter wells; polystyrene latex, in beads or microtiter plates; polyvinylidine fluoride; diazotized paper; nylon membranes; activated beads, and the like.

Typically, the solid support is first reacted with the biological sample (or the gp120 proteins), washed and then the antibodies, (or a sample suspected of containing antibodies), applied. After washing to remove any non-bound ligand, a secondary binder moiety is added under suitable binding conditions, such that the secondary binder is capable of associating selectively with the bound ligand. The presence of the secondary binder can then be detected using techniques well known in the art. Typically, the secondary binder will comprise an antibody directed against the antibody ligands. A number of anti-human immunoglobulin

(Ig) molecules are known in the art (e.g., commercially available goat anti-human Ig or rabbit anti-human Ig). Ig molecules for use herein will preferably be of the IgG or IgA type, however, IgM may also be appropriate in some instances. The Ig molecules can be readily conjugated to a detectable enzyme label, such as horseradish peroxidase, glucose oxidase, Beta-galactosidase, alkaline phosphatase and urease, among others, using methods known to those of skill in the art. An appropriate enzyme substrate is then used to generate a detectable signal.

5

10

15

20

25

30

Alternatively, a "two antibody sandwich" assay can be used to detect the proteins of the present invention. In this technique, the solid support is reacted first with one or more of the antibodies directed against Env (e.g., gp120), washed and then exposed to the test sample. Antibodies are again added and the reaction visualized using either a direct color reaction or using a labeled second antibody, such as an anti-immunoglobulin labeled with horseradish peroxidase, alkaline phosphatase or urease.

Assays can also be conducted in solution, such that the viral proteins and antibodies thereto form complexes under precipitating conditions. The precipitated complexes can then be separated from the test sample, for example, by centrifugation. The reaction mixture can be analyzed to determine the presence or absence of antibody-antigen complexes using any of a number of standard methods, such as those immunodiagnostic methods described above.

The modified Env proteins, produced as described above, or antibodies to the proteins, can be provided in kits, with suitable instructions and other necessary reagents, in order to conduct immunoassays as described above. The kit can also contain, depending on the particular immunoassay used, suitable labels and other packaged reagents and materials (i.e. wash buffers and the like). Standard immunoassays, such as those described above, can be conducted using these kits.

The Env polypeptides and polynucleotides encoding the polypeptides can also be used in vaccine compositions, individually or in combination, in e.g., prophylactic (i.e., to prevent infection) or therapeutic (to treat HIV following infection) vaccines. The vaccines can comprise mixtures of one or more of the modified Env proteins (or nucleotide sequences encoding the proteins), such as Env (e.g., gp120) proteins derived from more than one viral isolate. The vaccine may also be administered in conjunction with other antigens and immunoregulatory agents, for example, immunoglobulins, cytokines, lymphokines, and chemokines, including but not limited to IL-2, modified IL-2 (cys125-ser125), GM-CSF, IL-

12, γ-interferon, IP-10, MIP1β and RANTES. The vaccines may be administered as polypeptides or, alternatively, as naked nucleic acid vaccines (*e.g.*, DNA), using viral vectors (*e.g.*, retroviral vectors, adenoviral vectors, adeno-associated viral vectors) or non-viral vectors (*e.g.*, liposomes, particles coated with nucleic acid or protein). The vaccines may also comprise a mixture of protein and nucleic acid, which in turn may be delivered using the same or different vehicles. The vaccine may be given more than once (*e.g.*, a "prime" administration followed by one or more "boosts") to achieve the desired effects. The same composition can be administered as the prime and as the one or more boosts. Alternatively, different compositions can be used for priming and boosting.

5

10

15

20

25

30

The vaccines will generally include one or more "pharmaceutically acceptable excipients or vehicles" such as water, saline, glycerol, ethanol, etc. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present in such vehicles.

A carrier is optionally present which is a molecule that does not itself induce the production of antibodies harmful to the individual receiving the composition. Suitable carriers are typically large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycollic acids, polymeric amino acids, amino acid copolymers, lipid aggregates (such as oil droplets or liposomes), and inactive virus particles. Such carriers are well known to those of ordinary skill in the art. Furthermore, the Env polypeptide may be conjugated to a bacterial toxoid, such as toxoid from diphtheria, tetanus, cholera, etc.

Adjuvants may also be used to enhance the effectiveness of the vaccines. Such adjuvants include, but are not limited to: (1) aluminum salts (alum), such as aluminum hydroxide, aluminum phosphate, aluminum sulfate, etc.; (2) oil-in-water emulsion formulations (with or without other specific immunostimulating agents such as muramyl peptides (see below) or bacterial cell wall components), such as for example (a) MF59 (International Publication No. WO 90/14837), containing 5% Squalene, 0.5% Tween 80, and 0.5% Span 85 (optionally containing various amounts of MTP-PE (see below), although not required) formulated into submicron particles using a microfluidizer such as Model 110Y microfluidizer (Microfluidics, Newton, MA), (b) SAF, containing 10% Squalane, 0.4% Tween 80, 5% pluronic-blocked polymer L121, and thr-MDP (see below) either microfluidized into a submicron emulsion or vortexed to generate a larger particle size

5

10

15

20

25

30

emulsion, and (c) Ribi™ adjuvant system (RAS), (Ribi Immunochem, Hamilton, MT) containing 2% Squalene, 0.2% Tween 80, and one or more bacterial cell wall components from the group consisting of monophosphorylipid A (MPL), trehalose dimycolate (TDM), and cell wall skeleton (CWS), preferably MPL + CWS (DetoxTM); (3) saponin adjuvants, such as Stimulon™ (Cambridge Bioscience, Worcester, MA) may be used or particle generated therefrom such as ISCOMs (immunostimulating complexes); (4) Complete Freunds Adjuvant (CFA) and Incomplete Freunds Adjuvant (IFA); (5) cytokines, such as interleukins (IL-1, IL-2, etc.), macrophage colony stimulating factor (M-CSF), tumor necrosis factor (TNF), etc.; (6) detoxified mutants of a bacterial ADP-ribosylating toxin such as a cholera toxin (CT), a pertussis toxin (PT), or an E. coli heat-labile toxin (LT), particularly LT-K63 (where lysine is substituted for the wild-type amino acid at position 63) LT-R72 (where arginine is substituted for the wild-type amino acid at position 72), CT-S109 (where serine is substituted for the wild-type amino acid at position 109), and PT-K9/G129 (where lysine is substituted for the wild-type amino acid at position 9 and glycine substituted at position 129) (see, e.g., International Publication Nos. W093/13202 and W092/19265); and (7) other substances that act as immunostimulating agents to enhance the effectiveness of the composition.

Muramyl peptides include, but are not limited to, N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP), N-acteyl-normuramyl-L-alanyl-D-isogluatme (nor-MDP), N-acetylmuramyl-L-alanyl-D-isogluatminyl-L-alanine-2-(l'-2'-dipalmitoyl-sn-glycero-3-huydroxyphosphoryloxy)-ethylamine (MTP-PE), etc.

Typically, the vaccine compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection may also be prepared. The preparation also may be emulsified or encapsulated in liposomes for enhanced adjuvant effect, as discussed above.

The vaccines will comprise a therapeutically effective amount of the modified Env proteins, or complexes of the proteins, or nucleotide sequences encoding the same, and any other of the above-mentioned components, as needed. By "therapeutically effective amount" is meant an amount of a modified Env (e.g., gp120) protein which will induce a protective immunological response in the uninfected, infected or unexposed individual to which it is administered. Such a response will generally result in the development in the subject of a secretory, cellular and/or antibody-mediated immune response to the vaccine. Usually, such

a response includes but is not limited to one or more of the following effects; the production of antibodies from any of the immunological classes, such as immunoglobulins A, D, E, G or M; the proliferation of B and T lymphocytes; the provision of activation, growth and differentiation signals to immunological cells; expansion of helper T cell, suppressor T cell, and/or cytotoxic T cell.

5

10

15

20

25

30

Preferably, the effective amount is sufficient to bring about treatment or prevention of disease symptoms. The exact amount necessary will vary depending on the subject being treated; the age and general condition of the individual to be treated; the capacity of the individual's immune system to synthesize antibodies; the degree of protection desired; the severity of the condition being treated; the particular Env polypeptide selected and its mode of administration, among other factors. An appropriate effective amount can be readily determined by one of skill in the art. A "therapeutically effective amount" will fall in a relatively broad range that can be determined through routine trials.

Once formulated, the nucleic acid vaccines may be accomplished with or without viral vectors, as described above, by injection using either a conventional syringe or a gene gun, such as the Accell® gene delivery system (PowderJect Technologies, Inc., Oxford, England). Delivery of DNA into cells of the epidermis is particularly preferred as this mode of administration provides access to skin-associated lymphoid cells and provides for a transient presence of DNA in the recipient. Both nucleic acids and/or peptides can be injected either subcutaneously, epidermally, intradermally, intramucosally such as nasally, rectally and vaginally, intraperitoneally, intravenously, orally or intramuscularly. Other modes of administration include oral and pulmonary administration, suppositories, needle-less injection, transcutaneous and transdermal applications. Dosage treatment may be a single dose schedule or a multiple dose schedule. Administration of nucleic acids may also be combined with administration of peptides or other substances.

While the invention has been described in conjunction with the preferred specific embodiments thereof, it is to be understood that the foregoing description as well as the examples which follow are intended to illustrate and not limit the scope of the invention. Other aspects, advantages and modifications within the scope of the invention will be apparent to those skilled in the art to which the invention pertains.

Experimental

5

10

15

20

25

30

Below are examples of specific embodiments for carrying out the present invention. The examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

Efforts have been made to ensure accuracy with respect to numbers used (e.g., amounts, temperatures, etc.), but some experimental error and deviation should, of course, be allowed for.

EXAMPLE 1

A.1. Best-Fit and Homology Searches

The crystal structure of HXB-2 gp 120 was downloaded from the Brookhaven database (COMPLEX (HIV ENVELOPE PROTEIN/CD4/FAB) 15-JUN-98 1GC1 TITLE: HIV-1 GP120 CORE COMPLEXED WITH CD4 AND A NEUTRALIZING HUMAN ANTIBODY). Beta strands 3, 2, 21, and 20 of gp 120 form a sheet near the CD4 binding site. Strands β -3 and β -2 are connected by the V1/V2 loop. Strands β -21 and β -20 are connected by another small loop. The H-bonds at the interface between strands β -2 and β -21 are the only connection between domains of the "lower" half of the protein (joining helix alpha 1 to the CD4 binding site). This beta sheet and these loops mask some antigens (e.g., antigens which may generate neutralizing antibodies) that are only exposed during the CD4 binding.

Constructs that remove enough of the beta sheet to expose the antigens in the CD4 binding site, but leave enough of the protein to allow correct folding were designed. Specifically targeted were modifications to the small loop and, optional deletion of the V1/V2 loops. Three different types of constructs were designed: (1) constructs encoding polypeptides that leave the number of residues making up the entire 4-strand beta sheet intact, but replace one or more residues; (2) constructs that encode polypeptide having at least one residue of at least one beta strand excised or (3) constructs encoding polypeptides having at least two residues of at least one beta strand excised. Thus, a total of 6 different turns were needed to rejoin the ends of the strands.

Initially, residues in the small loop (residues 427-430, relative to HXB-2) and connected beta strands (β -20 and β -21) were modified to contain Gly and Pro (common in beta turns). These sequences were then used as the target to match in each search. The

geometry of the target was matched to known proteins in the Brookhaven Protein Data Bank. In particular, 5-residue turns (including an overlapping single residue at the N-terminal, the 2 residue target turn and 2 overlapping residues at the C-terminal) were searched in the databases. In other words, these modified loops add a 2 residue turn that should be able to support a geometry that will maintain the beta-sheet structure of the wild type protein. The calculations were performed using the default parameters in the Loop Search feature of the Biopolymer module of the Syby1 molecular modeling package. In each case, the 25 best fits based on geometry alone were reviewed and, of those, several selected for homology and fit.

In addition, it was also determined what modifications could be made to remove most of the V1/V2 loop (residues 124-198, relative to HXB-2) yet leave the geometry of the protein intact. As with the small loop, constructs were also designed which excised one or more residues from the β -2 strand (residues 119-123 of HXB-2), the β -3 strand (residues 199-201 of HXB-2) or both β -2 and β -3. For these constructs, known loops were searched to match the geometry of a pentamer (including two remaining residues from the N-terminal side, a 2 residue turn and 1 C-terminal residue). For these searches, Gly-Gly was preferred as the insert along with at least one C-terminal substitution.

A.2. Small Loop Replacements

In one aspect, the native sequence was replaced with residues that expose the CD4 binding site, but leave the overall geometry of the protein relatively unchanged. For the small loop replacements, the target to match was: ASN425-MET426-GLY427-GLY428-GLY431. Results of the search are summarized in Table 1.

Table 1: Search of Small Loop (Asn425 through Gly431)

25

5

10

15

20

Rank	Sequence	RMSD	% Homology	Seq Id No.
Best fit	LYS-ASP-SER-ASN-ASN	0.16689	62.5	27
3	TYR-GLY-LEU-GLY-LEU	0.220308	62.5	28
4	GLU-ARG-GLU-ASP-GLY	0.241754	62.5	29
7	ARG-LYS-GLY-GLY-ASN	0.24881	100	30
12	TRP-THR-GLY-SER-TYR	0.26417	83.33	31

30

Based on these results, constructs encoding Gly-Gly (#7), Gly-Ser (#12) or Gly-Gly-Asn (#7) were recommended.

As V1/V2 and one or more residues of β -2 and β -3 are also optionally deleted in the modified polypeptides of the invention, known loops to match the geometry of the V1/V2 loop were also searched. The V1/V2 loop the target to match was: Lys121-Leu-122-Gly123-Gly124-Ser199. Some notable matches are shown in Table 2:

RMSD % Homology Seq Id. No. Rank Sequence GLN-VAL-HIS-ASP-GLU 68.75 32 0.154764 Best fit 2 0.15718 81.25 33 LYS-GLU-GLY-ASP-LYS 34 9 0.173731 68.75 ARG-SER-GLY-ARG-SER 0.175554 81.25 35 11 THR-LEU-GLY-ASN-SER 0.178772 93.75 36 16 HIS-PHE-GLY-ALA-GLY

Table 2: Search of V1/V2 loop (Lys121 through Ser199)

15

10

5

Based on these searches, constructs encoding Gly-Asn in place of V1/V2 were recommended.

A.3. One Additional Residue Excisions

20

For a slightly truncated small loop, one more residue was trimmed from each beta strand to slightly shorten the beta sheet. The target to match was: ILE424-ASN425-GLY426-GLY427-LYS432. Results are shown in Table 3:

Table 3: Search of Beta sheet shortened by One residue (Ile424 through Lys432)

25

Rank	Sequence	RMSD	% Homology	Seq Id No.
Best fit:	ARG-MET-ALA-PRO-VAL	0.316805	58.33	37
Best	ASP-SER-ASP-GLY-PRO	0.440896	83.33	38
hom:				

Although these searches showed more variation and worse fits than the previous truncation, the Pro-Val or Pro-Leu encoding constructs were very similar. Accordingly, Ala-Pro encoding constructs were recommended.

Sequences encoding gp120 polypeptides having V1/V2 deleted and an additional residue from β -2 or β -3 excised were also searched. The V1/V2 loop the target to match was: VAL120-LYS121-GLY122-GLY123-VAL200. Some notable matches are shown in Table 4.

Table 4: Search of V1/V2 loop (Val120 through Val200)

10	Rank	Sequence	RMSD	% Homology	Seq Id No
	Best fit:	THR-VAL-ASP-PRO-TYR	0.400892	58.33333	39
	2	SER-THR-ASN-PRO-LEU	0.402575	54.16667	40
	3	THR-ARG-SER-PRO-LEU	0.403965	58.33333	41
	7	ARG-MET-ALA-PRO-VAL	0.440118	58.33333	42

15

20

5

The construct encoding Ala-Pro (e.g., #7) was recommended.

A.4. Further Excisions

In yet another truncation, an additional residue was trimmed from the β -20 and β -21 strands to further shorten the beta sheet. The target to match was ILE423-ILE424-GLY425-GLY426-ALA433. Notable matches are shown in Table 5.

Table 5: Search of Beta sheet shortened by Two Residues (Ile423 through Ala433)

Rank	Sequence	RMSD	% Homology	Seq Id No
Best fit:	THR-TYR-GLU-GLY-VAL	0.130107	79.16666	43
2	GLN-VAL-GLY-ASN-THR	0.138245	79.16666	44
3:	THR-VAL-GLY-GLY-ILE	0.153362	100	45

25

A construct encoding Gly-Gly (e.g., #3), which has 100% homology, was recommended.

Also searched were sequences encoding a deleted V1/V2 region and at least two residues excised from β -2, β -3 or at least one residue excised from β -2 and β -3. The target to match was: CYS119-VAL120-GLY121-GLY122-ILE201. Notable matches are shown in Table 6.

5

Table 6: Search of V1/V2 loop (Cys119 through Ile201)

Rank	Sequence	RMSD	% Homology	Seq Id No
Best fit:	ASP-LEU-PRO-GLY-CYS	0.250501	75	46
4	ASP-VAL-GLY-GLY-LEU	0.290383	100	47

10

15

It was determined that both constructs would be used.

B.1. Constructs encoding modified Env polypeptides

As described above, the native loops extruding from the 4- β antiparallel-stands were excised and replaced with 1 to 3 residue turns. The loops were replaced so as to leave the entire β -strands or excised by trimming one or more amino acid from each side of the connected strands. The ends of the strands were rejoined with turns that preserve the same backbone geometry (*e.g.*, tertiary structure of β -20 and β -21), as determined by searching the Brookhaven Protein Data Bank.

Table 7A is a summary of the truncations of the variable regions 1 and 2 recommended for this study, as determined in Example 1.A. above.

Table 7A

V1/V2 Modifications	SEQ ID NO	Figure
-LEU122-GLY-ASN-SER199	7	10
-LYS121-ALA-PRO-VAL200-	6	9
-VAL120-GLY-GLY-ILE201-	4	7
-VAL120-PRO-GLY-ILE201B-	5	8
-VAL120-GLY-ALA-GLY-ALA204-	3	6
-VAL120-GLY-GLY-ALA-THR202-	8	11
-VAL127-GLY-ALA-GLY-ASN195-	25	28

10

15

5

As previously noted, the polypeptides encoded by the constructs of the present invention are numbered relative to HXB-2, but the particular amino acid residue of the polypeptides encoded by these exemplary constructs is based on SF-162. Thus, for example, although amino acid residue 195 in HXB-2 is a serine (S), constructs encoding polypeptides having then wild type SF162 sequence will have an asparagine (N) at this position. Table 7B shows just three of the variations in amino acid sequence between strains HXB-2 and SF162. The entire sequences, including differences in residue and amino acid number, of HXB-2 and SF162 are shown in the alignment of Figure 2 (SEQ ID NOs:1 and 2).

20

Table 7B

HXB-2 amino acid number	HXB-2 Residue	SF162 Residue/amino acid number		
128	Serine (S)	Thr (T)/114		
195	Serine (S)	Asn (N)/188		
426	Met (M)	Arg (R)/411		

25

30

Constructs containing deletions in the β -20 strand, β -21 stand and small loop were also constructed. Shown in Table 8 are constructs encoding truncations in these regions. The constructs in Table 8 are numbered relative to HXB-2 but the unmodified amino acid sequence is based on SF162. Thus, the construct encodes an arginine (Arg) as is found in

SF162 in the amino acid position numbered 426 relative to HXB-2 (See, also, Table 7B). Changes from wildtype (SF162) are shown in bold in Table 8B.

Table 8

Small Loop/β-20 and β-21 (Modified)	SEQ ID NO	Figure
-TRP427-GLY-GLY431-	9	12
-ARG426- GLY-GLY -GLY431-	10	13
-ARG426- GLY-SER -GLY431B-	11	14
-ARG426-GLY-GLY-ASN-LYS432-	12	15
-ASN425- ALA-PRO -LYS432-	13	16
-ILE424-GLY-GLY-ALA433-	14	17
-ILE423-GLY-GLY-MET434-	15	18
GLN422-GLY-GLY-TYR435-	16	19
-GLN422- ALA-PRO- TYR435B-	17	20

15

10

5

The deletion constructs shown in Tables 7 and 8 for each one of the β-strands and combinations of them are constructed. These deletions will be tested in the Env forms gp120, gp140 and gp160 from different HIV strains like subtype B strains (e.g., SF162, US4, SF2), subtype E strains (e.g., CM235) and subtype C strains (e.g., AF110968 or AF110975).

20 Exemplary constructs for SF162 are shown in the

Figures and are summarized in Table 9. As noted above in Figure 2 and Table 7B, in the bridging sheet region, the amino acid sequence of SF162 differs from HXB-2 in that the Met426 of HXB-2 is an Arg in SF162. In Table 9, V1/V2 refers to deletions in the V1/V2 region; # bsm refers to a modification in the bridging sheet small loop.

25

Table 9						
Construct	Seq. Id.	Fig.	Modification/Amino acid sequence			
Val120-Ala204	3	6	V1/V2: Val120-Gly-Ala-Gly-Ala204			
Val120-Ile201	4	7	V1/V2: Val120-Gly-Gly-Ile201			
Val120-Ile201B	5	8	V1/V2: Val120-Pro-Gly-Ile201			
Lys121-Val200	6	9	V1/V2: Lys121-Ala-Pro-Val200			

		Т	Cable 9
Construct	Seq. Id.	Fig.	Modification/Amino acid sequence
Leu122-Ser199	7	10	V1/V2: Leu122-Gly-Asn-Ser199
Val120-Thr202	8	11	V1/V2: Val120-Gly-Gly-Ala-Thr202
Trp427-Gly431	9	12	bsm: Trp427-Gly-Gly431
Arg426-Gly431	10	13	bsm: Arg426-Gly-Gly-Gly431
Arg426-Gly431B	11	14	bsm: Arg426-Gly-Ser-Gly431
Arg426-Lys432	12	15	bsm: Arg426-Gly-Gly-Asn-Lys432
Asn425-Lys432	13	16	bsm: Asn425-Ala-Pro-Lys432
Ile424-Ala433	14	17	bsm: Ile424-Gly-Gly-Ala433
Ile423-Met434	15	18	bsm: Ile423-Gly-Gly-Met434
Gln422-Tyr435	16	19	bsm: Gln422-Gly-Gly-Tyr435
Val127-Asn195	25	28	bsm: Val127-Gly-Ala-Gly-Asn195
Gln422-Tyr435B	17	20	bsm: Gln422- Ala-Pro -Tyr435
Leu122-Ser199; Arg426-Gly431	18	21	V1/V2/bsm: Leu122-Gly-Asn-Ser199 Arg4 Gly-Gly-Gly431
Leu122-Ser199; Arg426-Lys432	19	22	V1/V2/bsm: Leu122-Gly-Asn-Ser199 Arg4 Gly-Gly-Asn-Lys432
Leu122-Ser199-Trp427- Gly431	20	23	V1/V2/bsm: Leu122-Gly-Asn-Ser199 Trp42 Gly-Gly431
Lys121-Val200- Asn425-Lys432	21	24	V1/V2/bsm: Lys121-Ala-Pro-Val200 Asn4: Ala-Pro-Lys432
Val120-Ile201-Ile424- Ala433	22	25	V1/V2/bsm: Val120-Gly-Gly-Ile201 Ile424 Gly-Gly-Ala433
Val120-Ile201B-Ile424- Ala433	23	26	V1/V2/bsm: Val120-Pro-Gly-Ile201 Ile424 Gly-Gly-Ala43
Val120-Thr202; Ile424- Ala433	24	27	V1/V2/bsm: Val120-Gly-Gly-Ala-Thr202 Ile424-Gly-Gly-Ala433
Val127-Asn195; Arg426-Gly431	25	29	V1/V2/bsm: Val127-Gly-Ala-Gly-Asn195 Arg426-Gly-Gly-Gly431

Combinations of V1/V2 deletions and bridging sheet small loop modifications in addition to those specifically shown in Table 9 are also within the scope of the present invention. Various forms of the different embodiments of the invention, described herein, may be combined.

The first screening will be done after transient expression in COS-7, RD and/or 293 cells. The proteins that are expressed will be analyzed by immunoblot, ELISA, and for binding to mAbs directed to the CD4 binding site and other important epitopes on gp120 to determine integrity of structure. They will also be tested in a CD4 binding assay and, in addition, the binding of neutralizing antibodies, for example using patient sera or mAb 448D (directed to Glu370 and Tyr384, a region of the CD4 binding groove that is not altered by the deletions).

5

10

The immunogenicity of these novel Env glycoproteins will be tested in rodents and primates. The structures will be administered as DNA vaccines or adjuvanted protein vaccines or in combined modalities. The goal of these vaccinations will be to archive broadly reactive neutralizing antibody responses.

Claims:

What is claimed is:

5 1. A polynucleotide encoding a modified HIV Env polypeptide wherein the polypeptide has at least one amino acid deleted or replaced in the region corresponding to residues 420 to 436 relative to HXB-2 (SEQ ID NO:1).

- 2. The polynucleotide of claim 1, wherein the region corresponding to residues 124-10 198 relative to HXB-2 is deleted and at least one amino acid is deleted or replaced in the regions corresponding to the residues 119 to 123 and 199 to 210 relative to HXB-2 (SEQ ID NO:1).
- The polynucleotide of claim 1, wherein at least one amino acid in the region
 corresponding to residues 427 through 429 relative to HXB-2 (SEQ ID NO:1) is deleted or replaced.
- The polynucleotide of claim 2, wherein at least one amino acid of the in the region corresponding to residues 427 through 429 relative to HXB-2 (SEQ ID NO:1) is deleted or replaced.
 - 5. The polynucleotide of claim 1, wherein the amino acid sequence of the modified HIV Env polypeptide is based on strain SF162.
- 6. An immunogenic modified HIV Env polypeptide having at least one amino acid deleted or replaced in the region corresponding to residues 420 through 436, relative to HXB-2 (SEQ ID NO:1).
- 7. The polypeptide of claim 6, wherein one amino acid is deleted in the region corresponding to residues 420 through 436, relative to HXB-2 (SEQ ID NO:1).

8. The polypeptide of claim 6, wherein more than one amino acid is deleted in the region corresponding to residues 420 through 436, relative to HXB-2 (SEQ ID NO:1).

- 9. The polypeptide of claim 6, wherein at least one amino acid is replaced in the region corresponding to residues 420 through 436, relative to HXB-2 (SEQ ID NO:1).
 - 10. The polypeptide of claim 6, wherein at least one amino acid residue between about amino acid residue 427 and amino acid residue 429 relative to HXB-2 (SEQ ID NO:1) is deleted or replaced.

10

- 11. The polypeptide of claim 6, wherein the V1 and V2 regions of the polypeptide are truncated.
- 12. The polypeptide of claim 10, wherein the V1 and V2 regions of the polypeptideare truncated.
 - 13. The polypeptide of claim 6, wherein the amino acid sequence of the modified HIV Env polypeptide is based on strain SF162.
- 20 14. A construct comprising the nucleotide sequence depicted in Figure 6 (SEQ ID NO:3).
 - 15. A construct comprising the nucleotide sequence depicted in Figure 7 (SEQ ID NO:4).

- 16. A construct comprising the nucleotide sequence depicted in Figure 8 (SEQ ID NO:5).
- 17. A construct comprising the nucleotide sequence depicted in Figure 9 (SEQ ID30 NO:6).

18. A construct comprising the nucleotide sequence depicted in Figure 10 (SEQ ID NO:7).

- 19. A construct comprising the nucleotide sequence depicted in Figure 11 (SEQ IDNO:8).
 - 20. A construct comprising the nucleotide sequence depicted in Figure 12 (SEQ ID NO:9).
- 21. A construct comprising the nucleotide sequence depicted in Figure 13 (SEQ ID NO:10).
 - 22. A construct comprising the nucleotide sequence depicted in Figure 14 (SEQ ID NO:11).
 - 23. A construct comprising the nucleotide sequence depicted in Figure 15 (SEQ ID NO:12).
- 24. A construct comprising the nucleotide sequence depicted in Figure 16 (SEQ IDNO:13).

15

- 25. A construct comprising the nucleotide sequence depicted in Figure 17 (SEQ ID NO:14).
- 25 26. A construct comprising the nucleotide sequence depicted in Figure 18 (SEQ ID NO:15).
 - 27. A construct comprising the nucleotide sequence depicted in Figure 19 (SEQ ID NO:16).
 - 28. A construct comprising the nucleotide sequence depicted in Figure 20 (SEQ ID NO:17).

29. A construct comprising the nucleotide sequence depicted in Figure 21 (SEQ ID NO:18).

- 30. A construct comprising the nucleotide sequence depicted in Figure 22 (SEQ IDNO:19).
 - 31. A construct comprising the nucleotide sequence depicted in Figure 23 (SEQ ID NO:20).
- 32. A construct comprising the nucleotide sequence depicted in Figure 24 (SEQ ID NO:21).
 - 33. A construct comprising the nucleotide sequence depicted in Figure 25 (SEQ ID NO:22).
 - 34. A construct comprising the nucleotide sequence depicted in Figure 26 (SEQ ID NO:23).
- 35. A construct comprising the nucleotide sequence depicted in Figure 27 (SEQ ID NO:24).

15

- 36. A construct comprising the nucleotide sequence depicted in Figure 28 (SEQ ID NO:25).
- 37. A construct comprising the nucleotide sequence depicted in Figure 29 (SEQ ID NO:26).
 - 38. A vaccine composition comprising a polynucleotide encoding a modified Env polypeptide according to any one of claims 1-5.
 - 39. A vaccine composition comprising a polynucleotide construct encoding a modified Env polypeptide according to any of claims 14-37.

40. A vaccine composition comprising a modified Env polypeptide according to any of claims 6-13.

41. The vaccine composition of any of claims 38-40, further comprising an adjuvant.

5

- 42. A method of inducing an immune response in subject comprising, administering a polynucleotide according to any one of claims 1-5 in an amount sufficient to induce an immune response in the subject.
- 43. A method of inducing an immune response in subject comprising, administering a polynucleotide construct according to any one of claims 14-37 in an amount sufficient to induce an immune response in the subject.
- 44. A method of inducing an immune response in a subject comprising administering a composition comprising a modified Env polypeptide according to any one of claims 6-13, wherein the composition is administered in an amount sufficient to induce an immune response in the subject
- 45. The method of any of claims 42-44 further comprising administering an adjuvant to the subject.
 - 46. A method of inducing an immune response in a subject comprising
 - (a) administering a first composition comprising a polynucleotide according to any of claims 1-5 in a priming step and
 - (b) administering a second composition comprising a modified Env polypeptide according to any of claims 6-13, as a booster, in an amount sufficient to induce an immune response in the subject.
- 47. The method of claim 46 wherein the first composition or second composition further comprise an adjuvant.

48. The method of claim 46 wherein the first and second compositions further comprise an adjuvant.

gp120 core structure

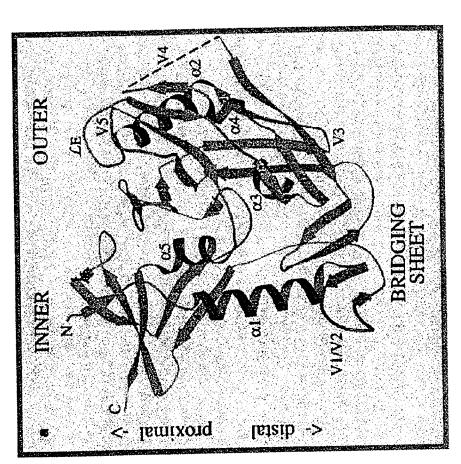


FIG. 1

WO 00/3	9303	2	1	65		T/US99/31272
HXB2 162 SF2 CM236 US4 Consensus	(1) (1) (1)	1 MRVKEKKQHLWRWEWMDAMKREL MKVKGTRRNKQHLWRWE- MRVKETOMNKPNLWKWEKHCQHLWRGE- MRVK YQHLWRWG	CCVLLL TLL ILL	CGAVFYSPS LGVLMIC-S LGVVIIC-S LGVLMIC-R	TEK VEKAN (CYE: TEKAN (CYE: Y SNN 900 (CYE)	ŵK WK WR WK
HXB2 162 SF2 CM236 US4 Consensus	(41) (46) (46) (41)	51 EAT WWW. CASDAKAYDTE EAT WWW. CASDAKAYDTE EAT WWW. CASDAKAYDTE DAD WWW. CASDAKAYDTE EAT WWW. CASDAKAYDTE EAT WWW. CASDAKAYDTE	V : 60024 (2) V : 6002 (2002 V : 6002 (2) A : 6002 (2)	o (Alexandra) Signatura Ministration (E)	OEVV (V M 220); POEVV G M 230; POEVV G MENIX SOEWHE MENIX	ISW ISW ISW
HXB2 162 SF2 CM236 US4 Consensus	(96) (96) (91)	KKIN KIN BOMO SA I I SERVITA KKIN KIN BOMO SA I I SERVITA KKIN KIN BOMH SA I I SERVITA	Stiff Stiff	HATVE :	DL DNL DL NAK DKLTGSTNGTNS	
HXB2 162 SF2 CM236 US4 Consensus	(129) (134) (135)	151KNDTNTNESSGKNETNTKESNWGKETNTNESNWLTNVNNITEVSN GTNSTSGTNETSTNETDS	KEMD-R KEET-K TIGNIT WEK Y PE	GEIK <mark>NESEK</mark> GEIK <mark>NESE</mark> NE GEIKNESENE GEIKNESENI	TSIRGKYÖKEYAF TTSIRN MOKEYAL TTSIRD MOKENAL TTERROKKOVUJAL	Y ER EY
HXB2 162 SF2 CM236 US4 Consensus	(171) (176) (179)	201 K.D. W. IDNDTTS K.DVV IDNDNTS N.DVV IDNASTITNYTN K. W. IDNKTSSE K.DVV IDNDNAS KLDVVPIDND TS	RIN RIH RIN RIN	STSWITSWIS SRSWITS STSWIKE STSWITSS	A CONTRACTOR OF THE CONTRACTOR	. ve - e - e - e - e
HXB2 162 SF2 CM236 US4 Consensus	(216) (226) (226) (236)	251 • FACCONNET CET PLET FACCONNET NEK 25T FACCONNET NEW 25T FACCO	1005 T	esuak uRe v Husteri <mark>R</mark> Subsk des v Juhak Revva	NOS ADEV GSS ADEV GSS ADEV NOS ADEV	VŽ VI VI
HXB2 162 SF2 CM236 US4 Consensus	(266) (276) (276) (286)	301 RSVNFED KORVOENT RSENFED VOKE RSENFEN VOKE RSENFEN VOENE RSENFEND VOENE RSENFEND VOENE RSENFEND VOENE	50/E 60(6 50/A 80(6 50/E 80(6 50/E 80(6	T <mark>eennine</mark> rks Teenninerks Teesinerts Teenninerks	REQREPGRAEVTI TE	D BR BD

FIG. 2A

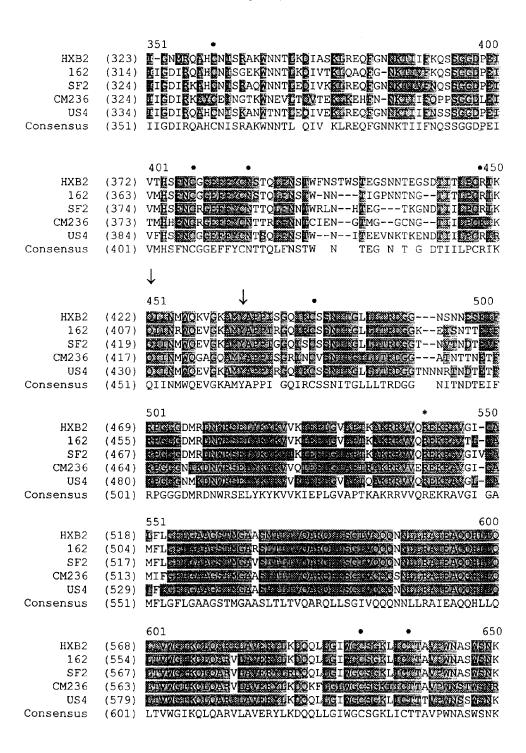


FIG. 2B

WO 00/39303			4	/	65			PCT/US9	9/31272
HXB2 162 SF2 CM236 US4 Consensus	(618) (604) (617) (613) (629) (651)	ELDOTONNMTOMEN ELE PENDAMTOMON EYE EENAAM CIEN ELT ZENDAMTOMEN	E (6) E (2) E (2) E (3)	DNYI DNYI SNYI GNYI	NLTY NTTY NQTYI GLTYI	LIEESON LEESON ENTESON NLIEI	OGEKNEQE GGEKNEQE GGEKNEQE	OGSERVA GGSERVA GGSERVA GGSERVA	
HXB2 162 SF2 CM236 US4 Consensus	(668) (654) (667) (663) (679) (701)	SAMMADS TN WERE	IKI IKI IKI	3810V 18851V 1881 I 1881 V	eefv eerv eer	SBRUVET RUVEA EA	egsje var even Gsje var erek Gsje var liever Gsjesijar ekster	orest Lef seest Lef leves Lef leves L	
HXB2 162 SF2 CM236 US4 Consensus	(713) (729)	OTRFPAPKG PRPE OTRLPVPKG 20KPD OFFHHOKE PORSE	G DE R DE G DE	2 602 2 602 6 602 2 603	RD ((b) RD ((b) QG ((b) RD ((c)	SP VH VR VD VR VS VNR VH	LuxtIVD FoxiIVS FoxiAVD LaxeIVD	MR4000 55 MR30 50 5 MR30 60 5	
HXB2 162 SF2 CM236 US4 Consensus		YR II DLI IAART MH. IF IAART MH. II IVARI	VELL VELL	R SH SRSS SR	 LKGLF	- KGNEAL - KGNEAL RIGHE GILV - KGWEALK	YWGN LOY YWWS ZOY YLGN LL Y	ISE KNS IESEKNS GUEDKIS SPEKSS	
HXB2 162 SF2 CM236 US4 Consensus	(813) (822)	VELFOIL AVAE VWLNET VAITE SLLDET VIVAG	G DE W 1008 G 1008	0 0 V. V 0 0 V. V 0 2 V.	AORIO AORA AOGA VORI	FL P B IL WH W IL P B B B B B B B B B B B B B B B B B B B	**************************************	A T 1 A	

FIG. 2C

(439) AAGTGCAACGACAAGAAGTTCAACGGCAGCGCCCCTGCA

Val120-Ile201

```
(439) AAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCA
Val120-Thr202
                (445) AAGTGCAACGACAAGAAGTTCAACGGCAGCGCCCCTGCA
Lvs121-Val200
                 (481) AAGTGCAACGACAAGAAGTTCAACGGCAGCGCCCCTGCA
     Consensus
                                                            560
                 (491) CCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCCC
Leu122-Ser199
                (521) CCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCCC
Val127-Asn195
                (479) CCAACGTGAGCACCGTGCAGCACCCACGGCATCCGCCC
Val120-Ile201B
                 (473) CCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCCC
 Val120-Ala204
                 (479) CCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCCC
Val120-Ile201
                 (479) CCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCCC
Val120-Thr202
                 (485) CCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCCC
Lys121-Val200
                 (521) CCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCCC
     Consensus
                       561
                 (531) CGTGGTGAGCACCCAGCTGCTGCAACGGCAGCCTGGCC
 Leu122-Ser199
                 (561) CGTGGTGAGCACCCAGCTGCTGCAACGGCAGCCTGGCC
 Val127-Asn195
                 (519) CGTGGTGAGCACCCAGCTGCTGAACGGCAGCCTGGCC
Val120-Ile201B
                 (513) CGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Val120-Ala204
                 (519) CGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Val120-Ile201
                 (519) CGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Val120-Thr202
                 (525) CGTGGTGAGCACCCAGCTGCTGAACGGCAGCCTGGCC
Lys121-Val200
                 (561) CGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCC
     Consensus
                                                            640
                 (571) GAGGAGGCGTGGTGATCCGCAGCGAGAACTTCACCGACA
Leu122-Ser199
                 (601) GAGGAGGCGTGGTGATCCGCAGCGAGAACTTCACCGACA
Val127-Asn195
                 (559) GAGGAGGCGTGGTGATCCGCAGCGAGAACTTCACCGACA
Val120-Ile201B
                 (553) GAGGAGGCGTGGTGATCCGCAGCGAGAACTTCACCGACA
 Val120-Ala204
                 (559) GAGGAGGCGTGGTGATCCGCAGCGAGAACTTCACCGACA
 Val120-Ile201
                 (559) GAGGAGGCGTGGTGATCCGCAGCGAGAACTTCACCGACA
Val120-Thr202
                 (565) GAGGAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACA
 Lys121-Val200
                 (601) GAGGAGGCGTGGTGATCCGCAGCGAGAACTTCACCGACA
     Consensus
                                                            680
                 (611) ACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGA
Leu122-Ser199
                 (641) ACGCCAAGACCATCATCGTGCAGCTGAAGGAGGAGCGTGGA
Val127-Asn195
                 (599) ACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGA
Val120-Ile201B
                 (593) ACGCCAAGACCATCATCGTGCAGCTGAAGGAGCGTGGA
 Val120-Ala204
                 (599) ACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGA
 Val120-Ile201
                 (599) ACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGA
 Val120-Thr202
                 (605) ACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGA
Lys121-Val200
                 (641) ACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGA
     Consensus
                 (651) GATCAACTGCACCGGCCCAACAACACCCCGCAAGAGC
Leu122-Ser199
                 (681) GATCAACTGCACCCGCCCAACAACAACACCCGCAAGAGC
Val127-Asn195
                 (639) GATCAACTGCACCCGCCCCAACAACACCCCGCAAGAGC
Val120-Ile201B
                 (633) GATCAACTGCACCCGCCCCAACAACAACACCCCGCAAGAGC
 Val120-Ala204
                 (639) GATCAACTGCACCCGCCCCAACAACAACACCCGCAAGAGC
 Val120-Ile201
                 (639) GATCAACTGCACCCGCCCCAACAACAACACCCGCAAGAGC
Val120-Thr202
                 (645) GATCAACTGCACCCGCCCCAACAACAACACCCCGCAAGAGC
 Lys121-Val200
                 (681) GATCAACTGCACCCGCCCCAACAACAACACCCGCAAGAGC
     Consensus
                       721
                 (691) ATCACCATCGGCCCGGCGCGCCTTCTACGCCACCGGCG
 Leu122-Ser199
                 (721) ATCACCATCGGCCCGGCCGCCCTTCTACGCCACCGGCG
 Val127-Asn195
                 (679) ATCACCATCGCCCCGGCCGCCCTTCTACGCCACCGGCG
Val120-Ile201B
                 (673) ATCACCATCGGCCCGGCCGCCCTTCTACGCCACCGGCG
 Val120-Ala204
                 (679) ATCACCATCGGCCCGGCCGCCCTTCTACGCCACCGGCG
 Val120-Ile201
                (679) ATCACCATCGGCCCGGCCGCCTTCTACGCCACCGGCG
Val120-Thr202
                 (685) ATCACCATCGGCCCGGCCGCCCTTCTACGCCACCGGCG
 Lys121-Val200
                 (721) ATCACCATCGGCCCGGCCGCCTTCTACGCCACCGGCG
     Consensus
```

```
(959) ACAACACCAACGGCACCATCACCCTGCCCTGCCGCATCAA
Val120-Ile201B
               (953) ACAACACCAACGGCACCATCACCCTGCCCTGCCGCATCAA
Val120-Ala204
                (959) ACAACACCAACGGCACCATCACCCTGCCCTGCCGCATCAA
Val120-Ile201
                (959) ACAACACCAACGGCACCATCACCCTGCCCTGCCGCATCAA
Val120-Thr202
                 (965) ACAACACCAACGGCACCATCACCCTGCCCTGCCGCATCAA
Lys121-Va1200
                (1001) ACAACACCAACGGCACCATCACCCTGCCCTGCCGCATCAA
     Consensus
                       1041
                (1011) GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
Leu122-Ser199
                (1041) GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
Val127-Asn195
Vall20-Ile201B
                (999) GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG.
                (993) GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
Val120-Ala204
                (999) GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
Val120-Ile201
                (999) GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
Val120-Thr202
               (1005) GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
Lys121-Va1200
                (1041) GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
     Consensus
                                                           1120
                       1081
                (1051) TACGCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
Leu122-Ser199
                (1081) TACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
Val127-Asn195
                (1039) TACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
Val120-Ile201B
                (1033) TACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
Val120-Ala204
               (1039) TACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
Val120-Ile201
               (1039) TACGCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
Val120-Thr202
               (1045) TACGCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
Lys121-Val200
               (1081) TACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
    Consensus
                                                           1160
                       1121
Leu122-Ser199 (1091) ACATCACCGGCCTGCTGCTGACCCGCGACGGCGACGGCAAGGA
               (1121) ACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGA
Val127-Asn195
               (1079) ACATCACCGGCCTGCTGACCCGCGACGGCGCAAGGA
Val120-Ile201B
               (1073) ACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGA
Val120-Ala204
                (1079) ACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGA
Val120-Ile201
                (1079) ACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGA
Val120-Thr202
                (1085) ACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGA
Lys121-Val200
               (1121) ACATCACCGGCCTGCTGCTGACCCGCGACGGCGCAAGGA
    Consensus
               (1131) GATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGC
Leu122-Ser199
               (1161) GATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGC
Val127-Asn195
               (1119) GATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGC
Val120-Ile201B
               (1113) GATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGC
Val120-Ala204
               (1119) GATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGCC
Val120-Ile201
               (1119) GATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGC
Val120-Thr202
               (1125) GATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGC
Lys121-Val200
                (1161) GATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGC
    Consensus
                       1201
               (1171) GACATGCGCGACAACTGGCGCGCGAGCTGTACAAGTACA
Leu122-Ser199
               (1201) GACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACA
Val127-Asn195
              (1159) GACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACA
Val120-Ile201B
               (1153) GACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACA
Val120-Ala204
Val120-Ile201 (1159) GACATGCGCGACAACTGGCGCGAGCTGTACAAGTACA
Vall20-Thr202 (1159) GACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACA
Lys121-Val200 (1165) GACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACA
              (1201) GACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACA
    Consensus
                                                           1280
                      1241
               (1211) AGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAA
Leu122-Ser199
               (1241) AGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAA
Val127-Asn195
               (1199) AGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAA
Val120-Ile201B
               (1193) AGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAA
Val120-Ala204
               (1199) AGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAA
Val120-Ile201
```

Consensus (1481) AGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGT

11 / 65

```
1521
 Leu122-Ser199 (1491) GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG
 Vall27-Asn195 (1521) GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG
Val120-Ile201B (1479) GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG
 Val120-Ala204 (1473) GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG
 Val120-Ile201 (1479) GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG
 Vall20-Thr202 (1479) GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG
               (1485) GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG
 Lys121-Val200
     Consensus
               (1521) GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG
                                                          1600
               (1531) GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG
 Leu122-Ser199
               (1561) GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG
 Val127-Asn195
               (1519) GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG
Val120-Ile201B
               (1513) GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG
 Val120-Ala204
               (1519) GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG
 Val120-Ile201
               (1519) GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG
 Val120-Thr202
Lvs121-Val200
               (1525) GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG
               (1561) GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG
     Consensus
                                                         1640
               (1571) CCGTGCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGA
Leu122-Ser199
Val127-Asn195
               (1601) CCGTGCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGA
Val120-Ile201B
               (1559) CCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGA
 Val120-Ala204
               (1553) CCGTGCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGA
               (1559) CCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGA
Val120-Ile201
               (1559) CCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGA
Val120-Thr202
Lys121-Val200 (1565) CCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGA
     Consensus (1601) CCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGA
Leu122-Ser199
              (1611) CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGC
Val127-Asn195 (1641) CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGC
Val120-Ile201B (1599) CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGC
               (1593) CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGC
Val120-Ala204
               (1599) CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGC
Val120-Ile201
               (1599) CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGC
Val120-Thr202
               (1605) CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGC
Lys121-Val200
               (1641) CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGC
    Consensus
                      1681
               (1651) GAGATCGACAACTACACCAACCTGATCTACACCCTGATCG
Leu122-Ser199
Val127-Asn195
              (1681) GAGATCGACAACTACACCAACCTGATCTACACCCTGATCG
              (1639) GAGATCGACAACTACACCAACCTGATCTACACCCTGATCG
Val120-Ile201B
              (1633) GAGATEGACAACTACACCAACCTGATCTACACCCTGATCG
Val120-Ala204
               (1639) GAGATCGACAACTACACCAACCTGATCTACACCCTGATCG
Vall20-Ile201
               (1639) GAGATEGACAACTACACCAACCTGATETACACCCTGATEG
Val120-Thr202
Lys121-Val200 (1645) GAGATCGACAACTACACCAACCTGATCTACACCCTGATCG
              (1681) GAGATCGACAACTACACCAACCTGATCTACACCCTGATCG
    Consensus
                                                         1760
              (1691) AGGAGAGCCAGAACCAGCAGGAGAACGAGCAGGAGCT
Leu122-Ser199
               (1721) AGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCT
Val127-Asn195
               (1679) AGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCT
Val120-Ile201B
              (1673) AGGAGACCAGAACCAGCAGAAGAACGAGCAGGAGCT
Val120-Ala204
              (1679) AGGAGAGCCAGAACCAGCAGGAGAACGAGCAGCAGCAGCT
Vall20-Ile201
Val120-Thr202 (1679) AGGAGAGCCAGAACCAGCAGGAGAACGAGCAGGAGCT
(1721) AGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCT
    Consensus
                      1761
Leu122-Ser199 (1731) GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTC
Vall27-Asn195 (1761) GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTC
```

```
Val120-Ile201B (1719) GCTGGAGCTGGACAGTGGGCCAGCCTGTGGAACTGGTTC
 Val120-Ala204 (1713) GCTGGAGCTGGACAGTGGGCCAGCCTGTGGAACTGGTTC
 Val120-Ile201
                (1719) GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTC
 Val120-Thr202
                (1719) GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTC
 Lys121-Val200
                (1725) GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTC
     Consensus
                (1761) GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTC
 Leul22-Ser199
                (1771) GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCA
 Val127-Asn195
                (1801) GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCA
                (1759) GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCA
Val120-Ile201B
 Val120-Ala204
                (1753) GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCA
 Val120-Ile201
                (1759) GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCA
 Val120-Thr202
               (1759) GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCA
 Lys121-Val200
               (1765) GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCA
     Consensus (1801) GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCA
 Leu122-Ser199 (1811) TGATCGTGGCCCTGGTGGCCTGCGCATCGTGTTCAC
 Val127-Asn195 (1841) TGATCGTGGGCGGCCTGGTGGGCCTGCTGTTCAC
Vall20-Ile201B (1799) TGATCGTGGCCGCCTGGTGGGCCTGCGCATCGTGTTCAC
 Val120-Ala204
               (1793) TGATCGTGGGCGGCCTGGGGCCTGCGCATCGTGTTCAC
 Val120-Ile201
                (1799) TGATCGTGGGCGCCTGGTGGGCCTGCGCATCGTGTTCAC
 Val120-Thr202
                (1799) TGATCGTGGGCGGCCTGGTGGGCCTGCGCATCGTGTTCAC
 Lys121-Val200
                (1805) TGATCGTGGGCGGCCTGGTGGGCCTGCGCATCGTGTTCAC
               (1841) TGATCGTGGGCGGCCTGGTGGGCCTGCGCATCGTGTTCAC
     Consensus
                      1881
                                                         1920
 Leu122-Ser199
               (1851) CGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGC
 Val127-Asn195
               (1881) CGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGC
Val120-Ile201B
               (1839) CGTGCTGAGCATCGTGAACCGGGTGCGCCAGGGCTACAGC
 Val120-Ala204
               (1833) CGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGC
 Val120-Ile201
               (1839) CGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGC
 Val120-Thr202
               (1839) CGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGC
 Lys121-Val200
               (1845) CGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGC
     Consensus
               (1881) CGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGC
 Leu122-Ser199
              (1891) CCCCTGAGCTTCCAGACCCGCTTCCCGGCCCCCCGCGGCC
 Val127-Asn195
              (1921) CCCCTGAGCTTCCAGACCGGCTTCCCCGGCCCCCCGGGCC
              (1879) CCCCTGAGCT/CCAGACCGGCX/IGCCGGCCCCCGGGGGC
Val120-Ile201B
               (1873) CCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCCCCGCGCCC
Val120-Ala204
Val120-Ile201
              (1879) CCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCCGCGGCC
Vall20-Thr202 (1879) CCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCCGCGGCC
Lys121-Val200 (1885) CCCTGAGCTTCCAGACCGCTTCCCGCCCCCCGCGGCC
    Consensus (1921) CCCCTGAGCTTCCAGACCGGCTTCCCCGCCCCCCGCGGCCC
Val127-Asn195
               (1961) CCGACCGCCCGAGGGCATCGAGGAGGAGGGCGCGAGCG
Vall20-Ile201B
               (1919) CCGACCGCCCCGAGGGCATCGAGGAGGAGGGCGGCGAGCG
Val120-Ala204
               (1913) CCGACCGCCCGAGGGCATCGAGGAGGAGGGCGCGAGCG
Val120-Ile201
               (1919) CCGACCGCCCGAGGGCATCGAGGAGGAGGGCGCGAGCG
               (1919) CCGACCGCCCGAGGGCATCGAGGAGGAGGGCGCGAGCG
Val120-Thr202
Lys121-Val200
               (1925) CCGACCGCCCGAGGGCATCGAGGAGGAGGGCGGCGAGCG
    Consensus (1961) CCGACCGCCCCGAGGGCATCGAGGAGGAGGGCGGCGAGCG
                      2001
Leu122-Ser199
               (1971) CGACCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTG
Val127-Asn195
              (2001) CGACCGCGACCGCAGCAGCCCCTGGTGCACGGCCTGCTG
Val120-Ile201B (1959) CGACCGCGACCGCAGCCCCCTGGTGCACGCCTGCTG
Val120-Ala204
              (1953) CGACCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTG
Val120-Ile201
              (1959) CGACCGCGACCGCAGCAGCCCCTGGTGCACGGCCTGCTG
```

```
Val120-Thr202 (1959) CGACCGCGACCGCAGCAGCCCCTGGTGCACGCCTGCTG
 Lys121-Val200 (1965) CGACCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTG
                 (2001) CGACCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTG
     Consensus
                 (2011) GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
 Leu122-Ser199
 Val127-Asn195 (2041) GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
Val120-Ile201B (1999) GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
 Val120-Ala204 (1993) GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
                (1999) GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
 Val120-Ile201
                 (1999) GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
 Val120-Thr202
                 (2005) GCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
 Lys121-Val200
                 (2041) GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
     Consensus
                        2081
                                                              2120
                 (2051) GCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCG
 Leu122-Ser199
 Val127-Asn195
                (2081) GCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCG
                (2039) GCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCG
Val120-Ile201B
 Val120-Ala204
                (2033) GCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCG
 Val120-Ile201
                 (2039) GCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCG
                 (2039) GCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCG
 Val120-Thr202
                (2045) GCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCG
 Lys121-Val200
                (2081) GCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCG
     Consensus
                        2121
                (2091) CATCGTGGAGCTGCTGGGGCCGCGGGCTGGGAGGCCCTG
 Leu122-Ser199
 Val127-Asn195
                (2121) CATCGTGGAGCTGCTGGGCCGCCGGGCTGGGAGGCCCTG
                (2079) CATOSTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTG
Vall20-Ile201B
                (2073) CATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTG
 Val120-Ala204
                (2079) CATCGTGGAGCTGCTGGGCCCCCGCGGCTGGGAGGCCCTG
 Val120-Ile201
 Val120-Thr202
                (2079) CATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTG
                (2085) CATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTG
 Lys121-Val200
     Consensus (2121) CATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTG
                        2161
 Leu122-Ser199 (2131) AAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGC
                (2161) AAGTACTGGGGGAACCTGCTGCAGTACTGGATCCAGGAGC
 Val127-Asn195
               (2119) AAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGC
(2113) AAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGC
(2119) AAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGC
(2119) AAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGC
Val120-Ile201B
 Val120-Ala204
 Val120-Ile201
 Val120-Thr202
                (2125) AAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGC
Lys121-Val200
                (2161) AAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGC
     Consensus
                        2201
                       TGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCAT
Leu122-Ser199
                (2171)
                (2201) TGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCAT
Val127-Asn195
Val120-Ile201B
                (2159) TGAAGAACAGCGCCGTGAGCCTGTTCGACGCGATCGCCAT
Val120-Ala204
                (2153) TGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCAT
                       TGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCAT
Val120-Ile201
                (2159)
                       TGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCAT
Val120-Thr202
                (2159)
Lys121-Val200
                (2165)
                       TGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCAT
     Consensus
                (2201) TGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCAT
               (2211) CGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC
Leul22-Ser199
Val127-Asn195 (2241) CGCCGTGGCCGAGGGCACCGCATCATCGAGGTGGCC
Vall20-Ile201B (2199) CGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC
Val120-Ala204 (2193) CGCCGTGGCCGAGGCCACCGCATCATCGAGGTGGCC
Vall20-Ile201 (2199) CGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC
Val120-Thr202 (2199) CGCCGTGGCCGAGGGCACCGCATCATCGAGGTGGCC
Lys121-Val200 (2205) CGCCGTGCCGAGGCCACCGACCGATCATCGAGGTGGCC
    Consensus (2241) CGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC
```

14 / 65

		2281 2320
Leu122-Ser199	(2251)	CAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCA
Val127-Asn195	(2281)	CAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCA
Val120-Ile201B	(2239)	CAGCGCATCGGCGGCGCCTTCCTGCACATCCCCCGCCGCA
Val120-Ala204	(2233)	CAGCGCATCGGCCGCCCTTCCTGCACATCCCCCGCCGCA
Val120-Ile201	(2239)	CAGCGCATCGGCCGCCCTTCCTGCACATCCCCCGCCGCA
Val120-Thr202	(2239)	CAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCA
Lys121-Val200	(2245)	CAGCGCATCGGCCGCCCTTCCTGCACATCCCCCGCCGCA
Consensus	(2281)	CAGCGCATCGGCCGCCCTTCCTGCACATCCCCCGCCGCA
		2321 2360
Leu122-Ser199	(2291)	TCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAGCG
Val127-Asn195	(2321)	TCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG
Val120-Ile201B	(2279)	TCCGCCAGGCCTTCGAGCGCGCCCTGCTGTAACTCGAGCG
Val120-Ala204	(2273)	TCCGCCAGGCCTTCGAGCGCGCCCTGCTGTAACTCGAG
Val120-Ile201	(2279)	TCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG
Val120-Thr202	(2279)	TCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG
Lys121-Va1200	(2285)	TCCGCCAGGGCTTCGAGCGCCCCCTGCTGTAACTCGAGCG
Consensus	(2321)	TCCGCCAGGGCTTCGAGCGCCCCTGCTGTAACTCGAG
		2361
Leu122-Ser199	(2331)	TGCT
Val127-Asn195	(2359)	
Val120-Ile201B	(2319)	TGCT
Val120-Ala204	(2311)	
Val120-Ile201	(2317)	
Val120-Thr202	(2317)	
Lys121-Va1200	(2325)	TGCT
Consensus	(2361)	

WO 00/39303		1	15	1	65			CT/US99/31272
Ile424-Ala433	(1)		exective	wilelen vi	lleletiVA ver	vácnáchácha		
Trp427-Gly431	(1)						سيست سيد بالجداد السادر	and the second s
Gln422-Tyr435B	(1)	cr:V:Vivike/de						A CONTRACTOR OF THE CONTRACTOR
Arg426-Gly431	(1)	A						
Ile423-Met434	(1)	***************************************	The second second					Balanton.
Gln422-Tyr435	(1)	CHANAMINO CLO		444				PP-00-00-00-00-00-00-00-00-00-00-00-00-0
Arg426-Lys432	(1)	ettantaringe leic						PARAMETER STATE OF THE STATE OF
Arg426-Gly431B	(1)	HAAVETICE						
Asn425-Lys432	(1)	टरंडे¥३५१£1४० ट (€						
Consensus	(1)	GAATTCGC 41	CACCA	ATGGAI	GCAATGA	AGAGAGGG		CT BO
Ile424-Ala433	(41)	entenie ente	CTGC	ic prejite	GAGGAG	(e-vilvo)e-tierij	(SIGISIO)	AG
Trp427-Gly431	(41)	STOTICE	e'ree	lezdey l e	erter (ter	el egitti	Cicle co	(G
Gln422-Tyr435B	(41)	GTGTGCTG	en Ge	синенис	GAGCAG	CHICGITY	CGCCC	S G
Arg426-Gly431	(41)							
Ile423-Met434	(41)	Grerecke	CAG	CITC N	GANCAGI	CHEICEVIII	CGCCC	NG
Gln422-Tyr435	(41)	eviewed ve	(e)Heres	GIIGNO	(e):\c(e):\c\	CTTCGTT	GCCC	NG .
Arg426-Lys432	(41)							
Arg426-Gly431B	(41)	GIGIGCIG						
Asn425-Lys432	(41)	ETETETE						
Consensus	(41)	GTGTGCTG 81	CTGCI	rGTGTG	GAGCAGT	CTTCGTTI		AG 20
Ile424-Ala433	(81)	GCC TCC	ta\e7a\rts	(C) (C) (C	(eje: (ej (e)	evenach)	recees	re
Trp427-Gly431	(81)	CGCCCTGG	AGAA	CHEC	grafe co	GTGTACT)	ceece	ĪĞ
Gln422-Tyr435B	(81)	CECCUICE	ACAM	CTC II	ery (GAL)	GTGTACT A	⇔elele/e	ŢĞ.
Arg426-Gly431	(81)	eger gree						
Ile423-Met434	(81)	Approximately designation of the contract of t						
Gln422-Tyr435	(81)	elele elentele						
Arg426-Lys432	(81)	OG V OKUNGIG						
Arg426-Gly431B		SECURG						
Asn425-Lys432	(81)		The state of the s					
Consensus	(81)		AGAAG	SCTGTG	GGTGACC	GTGTACTA		
T1 404 T1 400		121	-	enter monte promoti		100750000000000000000000000000000000000		50
Ile424-Ala433	(121)	ejelejeji ejilje						
Trp427-Gly431	(121)	CCDMGTS						
Gln422-Tyr435B	(121)	ale/clossressre	COCCUS INVESTMENTS	OCCUPATION OF STREET	***************************************			
Arg426-Gly431 Ile423-Met434	(121) (121)	del terre de deles care						2/1 57
Gln422-Tyr435	(121)	ekcjole/illeske						
Arg426-Lys432	(121)	e/c/e/e/e/e e						807 87
Arg426-Gly431B	(121)	Dar Haller College Committee College						
Asn425-Lys432	(121)	er (elemente	Carea semi-tempera	demonstration of the	CONTRACTOR SON CONTRACTOR	**************************************	Statement was a state	222.0
Consensus	(121)							
00110011040	(121)	161	0.2.00			.00010110		00
Ile424-Ala433	(161)	eleptelele	AAGG6	/eul#it@le)4(\$/4(\$(\$ (\$) 4	(ele)welezytez	v:\0/e\/(c	Ğ
Trp427-Gly431		ejejezwejejeje						
Gln422-Tyr435B	(161)	SCOACECC	eveyerek	(portette/ c	Y(e)\(\)(e(e)\(\)	(6(0)1/6(0)2(0))	едеўсуке,	Ğ
Arg426-Gly431		GEGACEC						
Ile423-Met434		ଞ୍ଜିତ୍ୟ ହୋଇଥିଲ						
Gln422-Tyr435		GCCACGOC						The state of the s
Arg426-Lys432		G(6)547/6/6/6/6(6						
Arg426-Gly431B		GC/C/A/C/C/C/C						
Asn425-Lys432		eleletatelele						
Consensus	(161)		AAGGC	CTACG	ACACCGA	GGTGCACA		
** 404 ** 155		201		6/11/2007 EMPL		The state of the s		10
Ile424-Ala433	(201)	clesses//e.a.c	ACCCC	Mecci	REPORTER	Carrago arcay)	(ଜନ୍ମର୍ଟ)	<u>1</u> G

FIG. 4A

WO 00/39303			16	1	65		PC'
Trp427-Gly431	(201)	egga.codo		•		erotterejerottete	(ક્ષણ ોલ
Gln422-Tyr435B	(201)					oterizateleletikiza	
Arg426-Gly431	(201)					(e,c47/e)e(e,(e37677)	
Ile423-Met434	(201)					[e]::}{e[e]e[e]:}\/\{e	
Gln422-Tyr435	(201)					s(e)1/c c(s(e)141/c	
Arg426-Lys432	(201)	eleteron de le	Neteleles	16 2 6 111	10000	cler:reiclelet.v. v	Deceme
Arg426-Gly431B	(201)					crevizere/presivite	
Asn425-Lys432	(201)			Manual Street,		0(674)0(0(6(0)2)A\(Phonon Company of the
Consensus	(201)					CGACCCCAAC	
Conscisus	(201)	241	ACGCC1	GCGIC	JCCCAC	COACCCCAAC	280
Ile424-Ala433	(241)		TOWN TOWN	e anno a co	Sala Para Sa	ena(enava(cuivuo):	
Trp427-Gly431	(241)					CACAVACIDADO	
Gln422-Tyr435B	(241)	ne chill Neu	v Velebr		none i	era(erava)e::#Wera	NO KATE III
Arg426-Gly431	(241)						
Ile423-Met434	(241)					GETEATA ENPOY	
						eweweenson eweweenson	AL HILL
Gln422-Tyr435	(241)						
Arg426-Lys432	(241)					estestationniest	
Arg426-Gly431B	(241)					GAGAACTITCA	
Asn425-Lys432	(241)					GAGEACTUCA	
Consensus	(241)		3CTGGA	GAACC	TGACC	GAGAACTTCA	
Ile424-Ala433	(281)	281	Neist engilie		767532 PAR	######################################	320
Trp427-Gly431	(281)					cic recovers	
_ _							
Gln422-Tyr435B	(281) (281)					1563 0 H 7669	
Arg426-Gly431 Ile423-Met434	(281)				NOCHUR	ile : : et (e)et. ile e : ejet (ejet.)	CHILAI
Gln422-Tyr435	(281)						
Arg426-Lys432	(281)					CCAVGAGGA	
Arg426-Gly431B	(281)					CCAVCAGE CCACGAGGA	
Asn425-Lys432	(281)	Eliter Harrison Indiana and Anna Anna Anna Anna Anna Anna An				Activities and an activities and a second activities activ	MINISTER PROPERTY.
Consensus	(281)					I SCACGAGGA TGCACGAGGA	
Consensus	(201)	321	MCAIG	GIGGE	IGCAGA	IGCACGAGGA	360
Ile424-Ala433	(321)					ere/v/erige/styce	
Trp427-Gly431	(321)						
Gln422-Tyr435B	(321)	Control of the last				ଞ୍ଜ : ୯ ୧୯ : ୧୭୧ : ୧	
Arg426-Gly431	(321)	Market Company	el as assess			ele elektrik	
Ile423-Met434							
Gln422-Tyr435	(321) (321)					code entej ej re code e <mark>delej</mark> e	
Arg426-Lys432	(321)					e e(e)e)il(e(e)es)ic	
Arg426-Gly431B	(321)	#15141K93:3040K			CONTRACTOR DESCRIPTION	co origeens	-
Asn425-Lys432	(321)					eje: e/il/eo/e (e	
Consensus	(321)					GCCCTGCGTG	
consensus	(321)	361	JOGACC	nonoc	CIGAA	GCCC1GCG1G	400
Ile424-Ala433	(361)		Vide(Nevi		re e	NC(Entelements	
Trp427-Gly431	(361)	A STATE OF THE PROPERTY OF THE PARTY OF THE	Contribution in Francisco Contribution in Cont			3.00(632 9 10.00	era de bosconomico (n. 12):
Gln422-Tyr435B	(361)	No let of other state.		20 V 10 C	1	zelos; wleusiko e	Tellie Vari
Arg426-Gly431	(361)					ue(e);\~e*;*;;*	
Ile423-Met434	(361)	No en el el el	, alakiri		uv ekire.	eleter (ervice	V #-21V1V #-21
Gln422-Tyr435	(361)					deaceard	
Arg426-Lys432	(361)		*****			verenteren atter	
Arg426-Gly431B	(361)	Need Com	211/212720			GOAGBAALG	MARIA DE LA
Asn425-Lys432	(361)					GCACCAACC	
Consensus	(361)	MANAGEM AND AND PROPERTY.		**********		TGCACCAACC	****
Consensus	(201)	401		U11000	LOCAC	CONCORNO	440
Ile424-Ala433	(401)		in are	AACAC	CACCA	a(emrejejava(ejeja)	
Trp427-Gly431	(401)	processor and the second secon	non-resident announce before	CONTRACTOR CONTRACTOR		CREE AGE	***************************************
Gln422-Tyr435B	(401)	Commenter and Co				verbeidiy.veled	MARKET PROPERTY AND ADDRESS OF THE PERSON NAMED AND ADDRESS OF
	(101)		A DA SAME		and the state of		

FIG. 4B

WO 00/39303		17 / 65	PCT/US99/31272
Arg426-Gly431	(401)	itelepologiales.v/toritelefiriteliteleficieliteleficielitelefi	VNECHICANNEC
Ile423-Met434		ોલાવાના સંસ્થાલ કરાયા છે. તેલા કરાયા કરવા કરવા છે. માટે કરાયા કર્યા છે. કરાયા કર્યા છે. કરાયા કર્યા છે. કર્યા ક	
Gln422-Tyr435		ત્રબદાઇએપ્રેલએપ્રેને જેવા કેમ્પ્રેલે કેમ્પ્યોલે કેમ્પ્રેલે કેમ્પ્યોલે કેમ્પ્સેલે કેમ્પ્સેલે કેમ્પ્રેલે કેમ્પ્રેલે કેમ્પ્સેલે કેમ્પ્સેલે કેમ્પ્સેલે કેમ્પ્	
Arg426-Lys432	(401)		
Arg426-Gly431B	(401)		
Asn425-Lys432	(401)	Netcicles/Actes/Attes/Ac	
Consensus	(401)		
		441	480
Ile424-Ala433	(441)	stercletetetetetetetetetetetetetetetetetetet	(c;c) :e);(€@:(\:@
Trp427-Gly431	(441)	zlojajajajajajajymentejajymentejajymikaj	lee (e)/c/4/4/6 -
Gln422-Tyr435B	(441)	અલદીવધાલાકાલા/લે?ક્ષીલાકાલા/કારાઓવાલા/લા/લાફોલોક	Ner (et:Neet:No.
Arg426-Gly431	(441)	elein nichten anwirtet atteit Getzele mitetist	to estado najo
Ile423-Met434	(441)	e.ele e.c.c.yelikhedirile. Alehiyele ilektirkelik	
Gln422-Tyr435	(441)	c.c.elololololotatatatainkoatatagitatakaholoitakoloitiikeatat	<i>(લક્ષાલ)લંબં</i>
Arg426-Lys432	(441)	era com concessos estantes protesta con estado en finificara.	
Arg426-Gly431B	(441)	बहुद्दिद्दिद्द्विद्दिर्द्दिक्ष्रिया १८८१ मा १००५ में १००५ १८०५ १८०५ १८०५ १८०५ १८०५ १८०५ १८०५	(c)c/((c))(v)E): @
Asn425-Lys432	(441)	ૺઌ૽ૡૺૡૡૡઌ૽ઌ૽ૡ૽ૹૡઌઌ૽૽ૡ૽૽૱ઌઌૡ૽ઌૡઌ૽ૡ૽૽૱	(clean/cravers) of
Consensus	(441)	CCGCGGCGAGATCAAGAACTGCAGCTTCAA	AGGTGACCACC
T1-404 71-400	(401)	481	520
Ile424-Ala433	(481)		04/070000000000000000000000000000000000
Trp427-Gly431	(481)	. (c/o): \. \. \. \. \. \. \. \. \. \. \. \. \.	MANAGEMENT STATES OF THE STATE
Gln422-Tyr435B	(481)	gereate no relativitativiteste celes testative leative he ta	
Arg426-Gly431	(481)	14cles145els1cles141/e9747/e374feles14s34786le31/e4147/e	
Ile423-Met434	(481)	;(e(e):'e(e):'\.'\.'\.'\.'\.'\.'\.'\.'\.'\.'\.'\.'	
Gln422-Tyr435	(481)	. (c	
Arg426-Lys432	(481)	$\mathcal{M}_{\mathcal{C}}(\mathcal{C}_{\mathcal{C}}(\mathcal{C})) = \mathcal{C}(\mathcal{C}(\mathcal{C}), \mathcal{V}_{\mathcal{C}}(\mathcal{C}), \mathcal{V}_{\mathcal{C}}(\mathcal{C}), \mathcal{V}_{\mathcal{C}}(\mathcal{C}(\mathcal{C}), \mathcal{C}_{\mathcal{C}}(\mathcal{C}), \mathcal{C}_{\mathcal{C}}(\mathcal$	MATA - 1990 CONTROL OF THE STATE OF THE STAT
Arg426-Gly431B	(481)	$\{(e, e, g, g, e, e, g, e, g, g,$	
Asn425-Lys432	(481)	ACCUMACION CONTROL CON	
Consensus	(481)	AGCATCCGCAACAAGATGCAGAAGGAGTAC 521	560
Ile424-Ala433	(521)	TWIN COURTERN WILLER HERMET, HERMET, HERMET, HAR	
Trp427-Gly431	(521)	Melyach ideax entelenteer etz hidrogazet	TO ENGLISH TO SE
Gln422-Tyr435B	(521)	Addyter midera centelestable et an 12) tratistice	7. 02. V N P. N P. N P. N C.
Arg426-Gly431	(521)	Mesty the considerations relegions by with the feature forces	hadava artiska (<mark>e</mark>
Ile423-Met434	(521)	specially and additional property of the statement	u na var sa koza e
Gln422-Tyr435	(521)	steint tiele, i steinterein ein hate ein thheist fartheis	nte netti cië
Arg426-Lys432	(521)	Activities in statule of elementary predebities element	n kanadang ke
Arg426-Gly431B	(521)	terrane en cleanerchiefen melek after sleaven) antere	
Asn425-Lys432	(521)	Activities selected and the selection of	New Notes (6) vie
Consensus	(521)	ACAAGCTGGACGTGGTGCCCATCGACAACG	
Ile424-Ala433	(561)	561 	600
Trp427-Gly431	(561)	chaire and rebut of the religions to the received	
Gln422-Tyr435B	(561)	July terrorite atternorment to rejective terrorite attendent	
Arg426-Gly431	(561)	instal sately resolutions also also states of the control of the	CONTRACTOR
Ile423-Met434	(561)	HINNEY THE CONTENTION OF THE C	
Gln422-Tyr435	(561)	and state constituted and object of the confidence of	
Arg426-Lys432	(561)	eventualistic ocalektarietatike stelestatieniselestike eleks	
Arg426-Gly431B	(561)	enni/zolai/cleanchinhesiai/enhelaitalesii/cleai/claien	
Asn425-Lys432	(561)	енки, (м. м.), с. јен бем бира (м. м. м	
Consensus	(561)	CTACAAGCTGATCAACTGCAACACCAGCGT	and the same of th
		601	640
Ile424-Ala433	(601)	દોભંજદાં ભુલાએ કરે જે દાનું ભૂત કરે છે. તે કર્યા છે છે. તે કર્યા છે	
Trp427-Gly431	(601)	eleichteicheichteichnetzteichniteletzteicheichnitei	erejestanojest, sim
Gln422-Tyr435B	(601)	cleen congressivites and the sand control of the sand	eratotrijet is toji
Arg426-Gly431	(601)	द्रेड्ः भरादेवलका रम्ददर्भवास्त ्रहासार्थका अद्यवस्ति स्वरूपार्	ल्लाकः, इत्राह्मकः, उत्तरीत
Ile423-Met434	(601)	संबक्षामुख्यस्थिकारः स्वयं हत्यः वृत्तः वृत्तः वृत्तः वृत्तः वृत्तः वृत्तः	

FIG. 4C

WO 00/39303		18 / 65	PCT/US99/31272
Gln422-Tyr435	(601)	ૡૺઌૺઌ૽ઌૺઌૡૺઌઌ૽૽ૹૢૡૡૺઌ૽ઌ૽૽૽ૺૡ૿ ૼૺઌ૽૽ઌ૽૽૽ૡૹૡૺૢૡઌૹૢૡઌ	elok veces ii
Arg426-Lys432	(601)		
Arg426-Gly431B	(601)	etele utetolokolatyyyetekiyesi, kelolaha etesikele lete uteletol	
Asn425-Lys432	(601)		
Consensus	(601)		and the same of th
	(/	641	680
Ile424-Ala433	(641)		
Trp427-Gly431		.(ଜ୍ୟାଟ୍ରମ୍ଡର୍ଡ୍ର୍ଲ୍ୟମ୍ବର୍ମ୍ବର୍ମ୍ବର୍ମ୍ବର୍ମ୍ବର୍ମ୍ବର୍ମ୍ବର୍ମ୍ବ	MANAGEMENT CONTRACTOR
Gln422-Tyr435B	(641)	୰ଢ଼୕୲ଽ୷୰ଢ଼୶ଢ଼ଢ଼ଢ଼ଢ଼୶ୡ୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷	
Arg426-Gly431	(641)		TO TAKE
Ile423-Met434	(641)		
Gln422-Tyr435	(641)		
Arg426-Lys432	(641)		
Arg426-Gly431B			
Asn425-Lys432	(641)		and the second s
Consensus		ACTGCGCCCCGCCGGCTTCGCCATCCTGAA	
combandas	(011)	681	720
Ile424-Ala433	(681)	MANACANATEAR AND MANACE CANTERNATE OF CONTRACTOR	
Trp427-Gly431	(681)	୭୯୬, ୧୯୬, ୧୯୭ <mark>,</mark> ୬୦୭ ପ୍ରତିଶ ଅଟଣ ପ୍ରତିଶ ହେଉଥିବା ଅଟ	
Gln422-Tyr435B		ङ । १.११ तम् ११ तम् । ११ तम् । ११ तम् । १९ तम्	
Arg426-Gly431		कार्यर्थात्रर प्रदेशक स्थापक स्थापक विकास कार्या कर्मा कार्या कर्मा कार्या कर्मा कर्मा कर्मा कर्मा कर्मा कर्मा	
Ile423-Met434	(681)	stivitel stront cotivite ele cittel clele elesantelotrote	DOMEO CHIMANO DE CAMORIO CAMOR
Gln422-Tyr435	(681)	W-10-10-10-10-10-10-10-10-10-10-10-10-10-	
Arg426-Lys432	(681)	arayyesia (Excidenty to ejelan) — delejela e o deletek	
Arg426-Gly431B	(681)		
Asn425-Lys432	(681)	BANGANGANGANA GEORNAGO NEORGANGANGA	
Consensus	(681)		new programment in the contract of the contrac
	, ,	721	760
Ile424-Ala433	(721)	TV COURS (CORS COS COS CERT) COS SESSES	CGTGAGCA
Trp427-Gly431	(721)	New Early Meanite experience of early	IGCTCACCA
Gln422-Tyr435B	(721)	: Novelet de los (elphones : o oles : o e (e) or e o eletetet	rélesses
Arg426-Gly431	(721)	AGOCHICGIACH COACHDAG GCCH I GUI GGGGE	rggzgag c a
Ile423-Met434	(721)	Act ichteletaetteloage van velos delejeletele	KGE/KE/AGE/A
Gln422-Tyr435	(721)	No electrical content (C) year con the circlestre conference (C)	UERSYCZAGC/A
Arg426-Lys432	(721)	to de la distrigaçõe/Acida de estra e elección ejec	ESCULPACION
Arg426-Gly431B	(721)		reenre cella
Asn425-Lys432	(721)	na oje sjemanjejske e na som najojeje toka si ir cole tokjejs	
Consensus	(721)	ACCGTGCAGTGCACCCACGGCATCCGCCCCG	rggtgagca
		761	800
Ile424-Ala433	(761)	decated in closic contentation of the content of th	
Trp427-Gly431	(761)	લું લેક્સ(લુંબ કદ પાકુલ પુરસ્કા હેલાલ દામાન દાલા મને દા લ ક ું લ ામ	est(cl∞eclett
Gln422-Tyr435B	(761)	ଜ୍ଜୁନ୍ନ ଅଟେ ଅଟେ ୧୯୬୬ ଅପରଅନ୍ତ ଓ ଅ <mark>ପର</mark> ସ୍ ର ୍ୟ	rc;, escopeu
Arg426-Gly431	(761)	siclosive in alternative elements of ends et Ar	(er/cje e e e zi
Ile423-Met434	(761)	S(c) = S(c) +	
Gln422-Tyr435	(761)	eres allegates enterior desires destructions elle elles elle	destruste (n.g.
Arg426-Lys432	(761)	ĸĸĸĸĸĊĸĸŶĸĸĸĸĸĸĸĸĸĸĸĸĸĸĸĸĸĸĸĸĸĸĸĸĸĸĸĸĸ	
Arg426-Gly431B	(761)	૦૦૦ : પર્વાનાયુદ્ધ કરવા હતા. પ્રાથમ કરવા હતા. (લાગ્રેન્ટ) (લાગ્રેન્ટ) (લાગ્રેન્ટ) (લાગ્રેન્ટ) (લાગ્રેન્ટ) (લાગ	- modern before minimum door
Asn425-Lys432	(761)	on a referre a factorie de la companie de la compa	
Consensus	(761)	CCCAGCTGCTGAACGGCAGCCTGGCCGAC	
T3 404 = 2 4 = 2		801	840
Ile424-Ala433	(801)	eleventi u ele cierte o el el el propio de la	
Trp427-Gly431	(801)	deserve administration of the property of the	
Gln422-Tyr435B		ciesticas regissionias enticatas electros estratorios	
Arg426-Gly431	(801)	ब्रोट ११६२ र १९८८ हो हो १९८८ हो १९८५ हो १९८५ हो १९८५ हो १९८५ हो १९८५	EIS GAVA GAS SE
Ile423-Met434	(801)	devictaviose eradedavera (elimbale i divelable)	
Gln422-Tyr435	(801)	elente.httoloe outeloettenningen ontget intrifol	
Arg426-Lys432	(801)	dencial control of the second states of the second	ର୍ଗ୍ୟ ହେଏ ହେଉଛି । ଅନ୍ୟାନ୍ତି

FIG. 4D

Arg426-Gly431B (801) CICHENTICECONCICENTAL PROPERTY CONTRACTOR AND Asn425-Lys432 (801) सन्।इतः,४५**०:अदार्थः,४६**।सनः,४६१५ प्रकान्। अपन्य सन्। इतः १४। । Consensus (801) GGTGATCCGCAGCGAGACTTCACCGACACGCCAAGACC Ile424-Ala433 (841)taketangetangetagetagetagatatetatetatetaletaketat/eitanstat/eitaletat Trp427-Gly431 (841)was missing the hold with the contract of the Gln422-Tyr435B .wife.tacoleagelest.clonkestatelest.catectex.clostestostatoratorae (841)Arg426-Gly431 (841) Ile423-Met434 (841) ANIGOVER CONTROLLENGE CONTROLLENGE AND ANIGOR Gln422-Tyr435 (841) Avervisioniscerrenteraceractresericerenteractreser Arg426-Lys432 Arg426-Gly431B (841) NIGAVOGI COAGGIRGAAGGAGAGAS SIRGGAGAGAGAGAGAGAGAGAGAGAGA Asn425-Lys432 (841) ANEMA COMECCIA CONTRACTORA CONTROL CONTR Consensus (841) ATCATCGTGCAGCTGAAGGAGCGTGGAGATCAACTGCA 881 920 Ile424-Ala433 (881) COURT OF STATE OF A TON SECTION SECTIO Trp427-Gly431 (881)elegeler (a.A.(e.)...a.).e.(e.).eeeeeeeee./e.)/e.)/ee: hisa(eee,/riceGe Gln422-Tyr435B (881)Arg426-Gly431 (881) PROCESS SAME TAGALOGO CONTRACTOR CONTR Ile423-Met434 (881) ## (CERT SEA NOTE FOR SEA SECTION OF SEA SECTION OF SECTION Gln422-Tyr435 (881) Xelekteria al-arterario all'intera deletaret ele compositione Arg426-Lys432 (881) MOSSOCIATACONACTAMENTOS CONTRACTOS MICROPALIACIO Arg426-Gly431B (881) Sendere de Anne Parel Parel Parel de Company de la Company de Company d Asn425-Lys432 (881) COLGOCOMARCIAN ARICAGON COLARGON CAGGAN CAGGAN CEG Consensus (881) CCCGCCCAACAACAACACCCGCAAGAGCATCACCATCGG 921 960 Ile424-Ala433 (921)ଅବସ୍ଥାତ ପ୍ରତି ଓ ଅନ୍ତର୍ଶ ଓ ଅନ୍ତ ୍ଟେସ-ସ୍ଥର୍ଗର୍ଟ୍ଟେସ୍ ଖଣ୍ଡ ଖଣ୍ଡ ଖଣ୍ଡ ଖଣ୍ଡ ଅନ୍ତର୍ମ ହେ । ଏହା ହେ । Trp427-Gly431 (921)Gln422-Tyr435B (921)Metalectical and the manufacture of the second of the seco Arg426-Gly431 ∞ଟିଟ୍ଲେଗମ ୬ ଅବସ୍ଥିତ । ଅଧିକ୍ର ଅଧିକ୍ର ଅଧିକ୍ର ଅଧିକ୍ର । (921)Ile423-Met434 (921) Preparete des la sur contra la ्राहित्य तम्रहालः ग्रेहायम्बद्धाः म्ये अतिहाससम्बद्धाः स्टब्स Gln422-Tyr435 (921)(921) ofcieloleje, loteje et e uniona, sejeje o re die eigleje et die eigleje et die eigleje et die eigleje eigen Arg426-Lys432 (921) Control of the Arg426-Gly431B Asn425-Lys432 (921) Color (clear color puller March Clear Geleic Activities 2017) Consensus (921) CCCCGGCCGCCCTTCTACGCCACCGGCGACATCATCGGC 961 Ile424-Ala433 Trp427-Gly431 (961) @/(6):44 / @/elelet/(6):16/6/1/1/(6):16/6/1/(6):1 Gln422-Tyr435B (961)ekiyosinga oʻshalariyo qidi oʻsho qaratigis vi yaki akaliyo qisha qiqis kileyiyeti oʻ (961) SACTING CERSONIC CICCON CURED NOT SERVED CERCENTAGENCIES Arg426-Gly431 atterning of cutted cotton relations in the legital actto with a Ile423-Met434 (961)Gln422-Tyr435 (961)chion remodely del close de cario de recepte enterater Arg426-Lys432 en skitke i delekte ardelo i tem odkeletivara e spatiskolejejejatatvakel (961)(961) Province of a circle for the circle and the circle for the c Arg426-Gly431B Asn425-Lys432 (961) GACAMA O SEACCE CAS CONSTRUCTOR CONTRACTOR Consensus (961) GACATCCGCCAGGCCCACTGCAACATCAGCGGCGAGAAGT 1040 Ile424-Ala433 (1001) ejerzykurzykerjkelete rejaktejerkerkykelet egyekratek keletatek Trp427-Gly431 (1001) EENALGIA CAAGGIGHETAACHAACHAACHAACHAACHACHACHACGAACHG (1001) Serva (MACCOE CHAPACHEAC) A VICE LICHACO A ACCOUNT CONTROL OF CONTROL Gln422-Tyr435B (1001) CETALETT OF A CHEMICAL CONTROL CONTRACTOR OF THE CONTROL CONTRO Arg426-Gly431 Ile423-Met434 (1001) CEANAGANGERCCEREANACION CANUSCRICAGEANACIONE E PACICIONAL CONTRA Gln422-Tyr435 (1001)eleginter interests to teliniteleginetiphelegicultativitelegieteleginele Arg426-Lys432 Arg426-Gly431B Asn425-Lys432

FIG. 4E

Consensus	(1001)	GGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGC
73 404 73 400		1041 1080
Ile424-Ala433	(1041)	eletricing readings to setting the setting of the s
Trp427-Gly431	(1041)	werderstandere :x:/en/y/en/gentere :
Gln422-Tyr435B	(1041)	ofert/enning efelv tatestat/est/oferthis encentrest/ent/ent/est/est/efe
Arg426-Gly431	(1041)	એએમના માત્રે છે. નાં લાન અંગણ જે પ્રકામ સ્ટામ સાંભળ છે. મેદી માત્ર હતા પ્રાપ્ત છે. પ્રદાસ પ્રદાસ માત્રે છે. મે
Ile423-Met434	(1041)	elocated in the electricated by the transfer of the strategies of the transfer electricated and the strategies of the st
Gln422-Tyr435	(1041)	સુંશકૃતિ , તાુના બેલા જાણકો અને ક્ષેત્ર કે મારે કર્યા છે. કાર્ય કર્યા માટે કર્યા છે. કાર્યકાર કર્યા છે કે કર્ય
Arg426-Lys432	(1041)	olechen niene eine habeton afterhassignafelen ein habet attelen bestelle siter
Arg426-Gly431B	(1041)	સાંસ્કૃતિકારિયાર કોઈલ્ડાસેર(સ્કૃત્રિસેર્સ્ટ) કરો કેલ્ડાસેર્સિક્ટ કેલ્ડા અને કેલ્ડાસેલ્ડાના સ્કૃતિ કોઈલ્ડાસેલ(સ
Asn425-Lys432	(1041)	એક્સ-૧૯- આ રાજ્યામાં એક્સ-૧૯-૧૬ (એ. ૧૯ -૧૯-૧૯) તાલુક છે. આ પ્રત્યામાં આ પ્રત્યામાં છે. તાલુક છે. તાલુક
Consensus	(1041)	CCAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAG
0000045	(1011)	1081 1120
Ile424-Ala433	(1081)	GG Sige of the safety of the safe safe side of the 44 effect of the safe side of the 44 effect of the 44 eff
Trp427-Gly431	(1081)	
		ૡૡ૽૽૽ૡૡ૽ૡૡ૽ૡ ૡ૽ઌઌ૽ઌઌ ઌૡ૽ૡૡઌઌઌઌઌઌઌઌ૽૽ૡ૽૱ઌઌઌઌઌઌઌઌઌઌઌઌૡૺૡૡૡ
Gln422-Tyr435B	(1081)	cicls electrical elect
Arg426-Gly431	(1081)	clefe(clefe) to possible for the free states and the state of the stat
Ile423-Met434	(1081)	deleteletete. You eletet fet fre fet med mig av transfer mer myterfes ele
Gln422-Tyr435	(1081)	er, derce her generation of the contraction of the contraction
Arg426-Lys432	(1081)	લં લુંલાલાં લાગા કુલ ૧૦૦૧/૧૯: ૧૦૯ મું છે. મુખ્યાન કુલા માનું અનુ છે. લ
Arg426-Gly431B	(1081)	હેલા (કુપલ પ્રાપ્ય લેલાવામાં માર્ગલો પ્રાપ્ય કરાય છે. જે મારા પ્રાપ્ય મારા પ્રાપ્ય મારા કરાય છે. જે મારા મારા મા
Asn425-Lys432	(1081)	द्रदार होत्य हो । । अब बोक्स कार्रमा बोक्स बेर्ग रहेत्त स्थान होता । स्वतंत्र स्थान होता ।
Consensus	(1081)	GGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG
		1121 1160
Ile424-Ala433	(1121)	encie ve tempo pomistante de la compania de la comp
Trp427-Gly431	(1121)	elected to the many releases to the releases of
Gln422-Tyr435B	(1121)	eleleladalustan kamina as magayya szeleny akatala sehmketky
Arg426-Gly431	(1121)	ભીંદાભાજ અંજ, (લો પંચ અમિલાલેક ૧૯૦ વર્ષ જ છે. ૧૯૩૧ અને ૧૦૦૦ અને ૧૧૦૦ અને મામાં દેવા છે.
Ile423-Met434	(1121)	eja eje u eliverim oja i orgineddeja ivitetite i atom at teordelindar. A
Gln422-Tyr435	(1121)	CIPICIEI (EPAC II) III II I
Arg426-Lys432	(1121)	ଅନ୍ୟ କ୍ରିଲ ଜଣ୍ଡ ଓଡ଼ିଆ ଓଡ଼ିଆ ଓଡ଼ିଆ ଓଡ଼ିଆ ହେଉଛି । ଅନ୍ୟୁଷ୍ଟ ଓଡ଼ି
Arg426-Gly431B	(1121)	es eleit satiezh kan hadiatkenaeta atuerkat ar mestre entekennearig
Asn425-Lys432	(1121)	esses enter in the series of t
Consensus	(1121)	GCGCCAGCTCTTCTACTGCAACAGCACCCAGCTGTTCAA
T) 101 T) 100		1161 1200
Ile424-Ala433	(1161)	gr. m. elepnez m. m. elepelez m. elepelez
Trp427-Gly431	(1161)	ante ante incienta introllidore incienta en la somitabilidade incienta en la somitabilidade inciente en la somitabilidade inciente en la somitabilidade en l
Gln422-Tyr435B	(1161)	करोट कर १५०० व्यवस्थातम् । संस्थिते । १८०० व्यवस्थातम् । वस्यापनार ।
Arg426-Gly431	(1161)	orac com a reletantemie, a statema eleje nem martina (elejen).
Ile423-Met434	(1161)	on isolopi, chiki me a mendektik kentik kentik ing kentik mendektik ing
Gln422-Tyr435	(1161)	on Actenia continua in proposito de la contrata de la contrata la contrata la contrata la contrata la contrata
Arg426-Lys432	(1161)	$\otimes_{\mathcal{C}}(C, x, t/r) = \otimes_{\mathcal{C}}(c, t/r) \otimes_{\mathcal{C}}(x, t/r)$
Arg426-Gly431B	(1161)	STACTOR CONTROL STATES A LANGE OF A COLONIA CONTROL OF A COLONIA CONTROL OF A COLONIA CONTROL OF A COLONIA
Asn425-Lys432	(1161)	અમાં પ્લામાન માટલામાં અપાસમાં આપણ અમેરી લાલા છે. જે કાર માટે પ્રાથમિક અને મા
Consensus	(1161)	CAGCACCTGGAACAACACCATCGGCCCCAACAACACCAAC
	. ,	1201 1240
Ile424-Ala433	(1201)	STATE OF THE PROPERTY OF THE P
Trp427-Gly431	(1201)	elegici destrita i concesso i Responsa i Malcian estra con informativi de la Malcia de Constanti de la responsa
Gln422-Tyr435B	(1201)	eselven interdecateodorite centrality dolly e
Arg426-Gly431	(1201)	elejestrane samitande kel eleste ja elest societates alterates al elegi
Ile423-Met434		
Gln422-Tyr435	(1201)	लेलाम् १८८८ । अ.२०६० मध्य एड १८८८ । १८५५ ४ व्यक्त १८६८ । १८५५ १८ व्यक्त १८६८ । १८५५ १८ व्यक्त १८६८ । १८५५ १८ व
_	(1201)	efeter (e.c.) - 1: er - (e.c. folker) efeter (e.k. folker) (c.
Arg426-Lys432	(1201)	ejcjestyczy, o tej foletej koje o to sielejcycie, s testy teje jycz skratoroj.
Arg426-Gly431B	(1201)	લું લુક્તું : (લાક), ધર્મ હતું, (લાક) કે પ્રકાર કે છે. જે જે લાક એ લોકો અને છે. જે કે જે જે જો છે. જો કે
Asn425-Lys432	(1201)	uclos/eles/anes/eleixances/eleixances/eleixances/eleixeleixi
Consensus	(1201)	GGCACCATCACCCTGCCCTGCCGCATCAAGCAGATCATCA
		1241 1280

FIG. 4F

21 / 65

Ile424-Ala433	(1240)	લેલબેલલ <u>ે</u> દે-લુક્ષેત્રમાં હયા મહાલા લેલ્કાલેલ લેક્સમાં હાલ <u>ે</u> ઉ
Trp427-Gly431	(1241)	AGCGCTGGEORGIO: PACIGIO VILC. PACIGIO VIC. COCCAMIO VIC
Gln422-Tyr435B	(1234)	GCGCC
Arg426-Gly431		
	(1241)	ः बुCGCGGC वसम्मिन् भागतिसम्बद्धाः ५४ सम्बद्धाः । वस्ति ।
Ile423-Met434	(1237)	Geologie And sylesistericity of the G
Gln422-Tyr435	(1234)	ଅଟ୍ରେଖ୍ୟର ୬/ହାଞ୍ଜ୍ୟର୍ଥ ହେଉଥାନ ନିର୍ମ୍ଦର
Arg426-Lys432	(1241)	-1@CGCGGCCGAAMAAACGCCCCCLUKCKK SCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
Arg426-Gly431B	(1241)	ACCGCGGCAECEGGAACEGGAACEAACCAACGGGGGGAACGG
Asn425-Lys432	(1241)	1/6
Consensus	(1241)	AC GGCGGCAAGGCCATGTACGCCCCCCCATCCG
00110011243	(12.11)	1281 1320
Ile424-Ala433	/12601	2020
	(1269)	celescontenteles melentacine en tinglication de la celes melentes
Trp427-Gly431	(1281)	ocklopacy soldkar oldbak olaan sola okko en se e
Gln422-Tyr435B	(1257)	ଵୄ୕୰ଽ୵୵୵ଽ <mark>୕୰୵ୡଽ୕ଌ୕ଵ୵</mark> ୶୵୵୵୵୵୷ଽ୵୵୷ଽ୵ଽ୵୷ୡ୷ୠ୷୷୷୷୷୷୷୷
Arg426-Gly431	(1281)	୶ୡୄ୕୵୰୰୰ୣ୰ୣୗ୰ୢ୲୴୲ଢ଼୴ଢ଼ଊ୳ଵୄ୶୷୷୰ୄ୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷
Ile423-Met434	(1263)	ଅଧିକ୍ରିୟ ବାୟକ୍ୟ ହେଉଥିଲି । ଏହି ବ୍ୟୁଷ୍ଟ ପ୍ରଥମ ଅଧିକ୍ରିୟ ହେଉଥିଲି । ଏହି ବ୍ୟୁଷ୍ଟ ପ୍ରଥମ ବ୍ୟୁଷ୍ଟ ବ୍ୟୁଷ୍ଟ ବ୍ୟୁଷ୍ଟ ବ୍ୟୁଷ୍ଟ
Gln422-Tyr435	(1257)	ଉପ୍ଟେଟ ଅ.୯୧: ୧୯ ସ୍ଲ୍ୟ ର ହେବା, ୧୭.୯୧୭: ୧୯.୯୭:୧୯.୯୭:୧୯.୯୭:୧୯.୯୭:୧୯
Arg426-Lys432	(1281)	श्रीत्वितिः ११९-१४१६। द्वेत्रविते १ ८ वर्गस्य ११, १००१ १५८ ११, १००१ १५८ ११, १००१ १५८ ११, १००१
Arg426-Gly431B	(1281)	ଜ୍ଞାମ୍ୟର୍ମ୍ୟ ହାମ୍ୟ ମଧ୍ୟ ହେଉଛି । ଜଣ୍ଡ ବ୍ୟବ୍ୟ ବ୍ୟବ୍ୟ କ୍ଷ୍ୟ ବ୍ୟବ୍ୟ କ୍ଷ୍ୟ କ୍ଷ୍ୟ ବ୍ୟବ୍ୟ କ୍ଷ୍ୟ କ୍ଷ୍ୟ କ୍ଷ୍ୟ କ୍ଷ୍ୟ କ୍ଷ
Asn425-Lys432	(1275)	ल्युद्रोक्षेत्र(सः प्रमण्डम् ह्राक्ष्मोद्राकः १८ ७ गृत्राहा रूपः १० गृह्यात् । ८० गृह्यात् । ८० ग्रह्यात् प्रह
Consensus	(1281)	CGGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTG
		1321 1360
Ile424-Ala433	(1309)	නේස: : ස්ප්රේදේවස්ස් හෝස්ස් (: 1678) වෙනු (ය. මේ. මේ.
Trp427-Gly431	(1321)	called the residence; (welcovered the filter also interpreted to each one
Gln422-Tyr435B	(1297)	००: इ.५१वर व्हानका महत्त्ववाच सम्बद्धाः इत्याप्त । सम्बद्धाः १००० का सम्बद्धाः । इत्याप्त । इत्याप्त । इत्याप
Arg426-Gly431	(1321)	www.ct.valencialct.cacjejejchon v.kejer. or.k. varienoj v. or ojer. or o
Ile423-Met434	(1303)	Michael Colombiace com an and alway way and a segme
Gln422-Tyr435		
	(1297)	exiter: (@@@@ch;(e'e)c @eje.c/ x.seicr.ser.u.@.tr. w.v.c/ , e.d/ oc. e
Arg426-Lys432	(1321)	୰୷୷୷୷୶ଢ଼୶୷୷ ୵୶୕୲ ୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷
Arg426-Gly431B	(1321)	क्षांतुद्धार/होदाल्ड लक्षार्थलालाहोलाहाः । ३.४४१० होर्गुलः ४४४० ग्रेस्टी होर्गुरः । ० छोत् । ० ह
Asn425-Lys432	(1315)	ଲ୍ଲାମ, (ଅଟ୍ରେମ୍ପ ଟ୍ୟକ୍ଟ ମ୍ବମନ୍ତ୍ର ଅନ୍ତର୍ଶ୍ୱର । । । । ଜ୍
Consensus	(1321)	CTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCG
		1361
Ile424-Ala433	(1349)	ACTION TO SELECTE SIGN CONTROL OF STATES
Trp427-Gly431	(1361)	detaile a tacie e e e e e e e e e e e e e e e e e e
Gln422-Tyr435B		
-	(1337)	ૡૡ૱ૺ૱૱ૡ૽ૡૡૡૡૡૡ૽૱૽ૡ૽૱ૡૡ૽ૡ૽ઌઌૡઌૡ૽ઌ૽૽ઌ૽૽ૡૡૡૡૡૡ૽૽૱૽૿૿ૢ
Arg426-Gly431	(1361)	મુંલામાં વિભાવસાય મુંબલા વાલાવાલા વાલાવાલા અમામ મુંબલા વાલામાં
Ile423-Met434	(1343)	૽૱૱ૢ૽ૡઌઌૡઌઌૣૡ૽ઌ૱૱ૡઌ૽૽૾૽ૡૡૡઌઌઌઌઌૡૡૡૡૡઌૢઌૢ૽ૺઌ૱ઌઌ ૺ
Gln422-Tyr435	(1337)	Activity with personal activities and the state of the second
Arg426-Lys432	(1361)	Activities of the detailed and the content of the c
Arg426-Gly431B	(1361)	्रदाः भागः । ग्रह्मावाः । श्रेष्टां स्टार्वे स्टार्वे । स्टार्थे स्टार्थे स्टार्थे स्टार्थे स्टार्थे स्टार्थे
Asn425-Lys432	(1355)	Terror seedester ender ver vielden virkere e
Consensus	(1361)	AGATCTTCCGCCCCGGCGGCGGCGACAACTG
Compensus	(1301)	
T1-424 B1-422	(1200)	1401 1440
Ile424-Ala433	(1389)	esetele treatest/capinethile ittleintmisture threfthleitetes apricite
Trp427-Gly431	(1401)	content yellow years you was and content years on the
Gln422-Tyr435B	(1377)	chalelectic alchitelections examplifying a color and selection and selection and selections and selections and selections are selected as a selection of the selection and selections are selected as a selection and selection are selected as a selected
Arg426-Gly431	(1401)	स्ट्रिल्लाक्षा व्यवस्था स्ट्रिल्ला स्ट्रिल्ला स्ट्रिल्ला स्ट्रिल्ला स्ट्रिल्ला स्ट्रिल्ला स्ट्रिल्ला स्ट्रिल्ला
Ile423-Met434	(1383)	eleleleste oleitelenettostatenatostaterihkelennetaitaitaithioisite
Gln422-Tyr435	(1377)	ૡૺૡૡૺ૱૱ૡૺૢ૱ૺૡૡૺૢૡૡૡ૽ૡ૱૱૱૱૱૱૱૱૱ૡૡૡૡૹ ૡ૱૱૱ ૹૡ <i>ૡ</i> ૱૱
Arg426-Lys432	(1401)	cjejcjej:/cjeje:/fcjeji/cji/jcji/s/cji/jcji/s/cjc/kejeskcji/s/ej.
Arg426-Gly431B	(1401)	
Asn425-Lys432		SOSSASSASSASSASSASSASSASSASSASSASSASSASS
	(1395)	geconematical contents and second contents and the
Consensus	(1401)	GCGCAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAG
***		1441 1480
Ile424-Ala433	(1429)	ૡઌ૽૽ૹૡૡઌૡૺૡૺૡૺૡઌૡૡ૽ૡ૽ૡઌઌ૽ઌ <i>૽૽૱</i> ૹ૽૽ૢૹ૽ૹ૽૽ૹ૽૽ૡૡઌ૱ૹ૽૽૱ૡૡ૽ૡૡઌઌૡૡૡ૽૽ઌૡૺ
Trp427-Gly431	(1441)	edepotential education (2006) en el estre el sustitute la estatut estate el estate estate el estate el estate

FIG. 4G

Gln422-Tyr435B	(1417)	કાલાં લામાના લાગ લાગ કામાં કામાં કામાં કામાં કામાં કામાં કામાં કામાં કામાં લાગ કામાં કામાં કામાં કામાં કામાં ક
Arg426-Gly431	(1441)	ୣ୵୶୲ୡଽ୕୷୷ୠ୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷
Ile423-Met434	(1423)	પારાં ભાગમાં એ લોલોલોલોલોલોલોલોલોલોલોલોલોલોલોલોલોલોલ
Gln422-Tyr435	(1417)	भविक अमहासाराज्यको स्वादिद्यादादादादादार १,१४८ हो स्वादिद्या । १,१४८ हो स्वादिद्या ।
Arg426-Lys432	(1441)	ાલાવામાં ભાષાનું લોકો કર્યા કાલા છે. તે કર્યા છે. તે કર્યા છે. તે કર્યા છે. તે કર્યા કર્યા કર્યા કર્યા કર્યા છે
Arg426-Gly431B	(1441)	्रक्षातः वेद्यदेशस्य । द्यंतिहेद्दरः द्रायम् । १९४४ व्यक्तिम् १९४५ व्यक्ति।
Asn425-Lys432	(1435)	ૹૺઌૢ૽ઌ૽ઌૢૢૢઌ૽૽ૡૺૡૺૡૺૡૺૡૺઌૼઌ૽૱ઌૡૺ૱૽ૺ૾૽૽ૹ૽૽ૹ૽૾ૺૺઌૺઌૺૹ૾ૹ૽૾ૹ૽૽ૺઌૺૡૺૡૺ <mark>ૡૺૡૺૡઌ</mark> ઌ૽૽૽૽ૡ
Consensus	(1441)	CCCCTGGGCGTGGCCCCCACCAAGGCCCAAGCGCCGCGTGG
	(2112)	1481 1520
Ile424-Ala433	(1469)	Relaycacionagis/sectetorecites/reprincedes/incisi
Trp427-Gly431	(1481)	HOD CONCENTRACED CONTRACED CHICAGO CONTRACED AND IN
Gln422-Tyr435B	(1457)	lacitavelet avolde entre control systematics el Vicolia
Arg426-Gly431	(1481)	re a companio Anguero de 1975 con tipo de 18
Ile423-Met434	(1463)	
Gln422-Tyr435	(1457)	ାପ୍ତ ଓ ପ୍ରତ୍ୟୁକ୍ତ ଏକ୍ଟେମ୍ବର ପ୍ରତ୍ୟୁକ୍ତ ପ୍ରତ୍ୟୁକ୍ତ ।
Arg426-Lys432		ଖଣ୍ଡ ୬/ଗ୍ରମ୍ମ ପ୍ରମୟ (ଗ୍ରହ୍ମ ଓ ଜଣ ଗ୍ରହ୍ମ ପ୍ରମୟ (ଜଣ ଅଧିକ୍ର ଜଣ କାଲ୍ଲା ପ୍ରମୟ ଆଧାର ୬/ଗ୍ରମ୍ମ ପ୍ରମୟ (ଗ୍ରହ୍ମ କର୍ମ କର୍ମ କର୍ମ କର୍ମ କର୍ମ କର୍ମ କର୍ମ କର
	(1481)	HE (directly engles with a color of the colo
Arg426-Gly431B	(1481)	Empto voletico a relación y con encelatio del elementario
Asn425-Lys432	(1475)	INCION CONTROCAVELA VARIO ELEGIO INCOMENDA EL MANTO INFO
Consensus	(1481)	TGCAGCGCGAGAAGCGCGCGTGACCCTGGGCGCCATGTT
T1-404 71 400	(1500)	1521 1560
Ile424-Ala433	(1509)	୕୶୳୲୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷
Trp427-Gly431	(1521)	acurdes a reporteduce que no realizadore
Gln422-Tyr435B	(1497)	୬ ସେ ଓ ଅପିରତ୍ୟୟ କ୍ଷତିଶ୍ୱରେ ହେ (କ୍ରେମ୍ବର ସହର ଓ ଓ
Arg426-Gly431	(1521)	: ୧୯.୧୯ ଅଟନ୍ତି କର୍ମ ଓ ଅନ୍ତର୍ଶ ମଧ୍ୟ ହେଉଛି । ଅନ୍ତର୍ଶ କର୍ମ ଅନ୍ତର୍ଣ କର୍ମ ଅନ୍ତର୍ଶ କର୍ମ କର୍ମ ଅନ୍ତର୍ଶ କର୍ମ ଅନ୍ତର୍ଶ କର୍ମ କର୍ମ ଅନ୍ତର୍ଶ କର୍ମ କର୍ମ ଅନ୍ତ
Ile423-Met434	(1503)	egylelecentromicolencelecoleta Weavelow Edesic e
Gln422-Tyr435	(1497)	ଜ୍ୟାନ୍ୟ ପର୍ପତ ଓ ଅନୁସାରଣ ବର୍ଷ ପ୍ରତ୍ୟୁ ପ୍ରତ୍ୟୁ ପ୍ରତ୍ୟୁ ପ୍ରତ୍ୟୁ ପ୍ରତ୍ୟୁ କର୍ଷ ଅନୁସାରଣ ହେଲି ।
Arg426-Lys432	(1521)	େ । ପ୍ରତି ଓ ପ୍ରତିକ୍ରଣ ବ୍ରତ୍ୟ କଥା । କଥା ବିଜୟ ପ୍ରତିକ୍ରଣ କଥା ।
Arg426-Gly431B	(1521)	େ ମାନ୍ତର ପ୍ରେମ୍ପର ଆଧାର କରିଥିଲି । ଏହି ବେ ଅଟଲି (ଲିଲ୍ମ୍କି (କରିଥିଲି କରିଥିଲି ।
Asn425-Lys432	(1515)	elevados machados de conserva en la serva
Consensus	(1521)	CCTGGGCTTCCTGGGCGCCGCCGGCAGCACCATGGGCGCC
#3 /0/ =3 / 00		1561 1600
Ile424-Ala433	(1549)	central e comatacementacement con certe certe
Trp427-Gly431	(1561)	genterie na energickanne na viterie ei
Gln422-Tyr435B	(1537)	a witer a cita in itely or enter yeter in a witer tee fert
Arg426-Gly431	(1561)	starticularing to be a particular to the contract and an entire resistance of the contract of
Ile423-Met434	(1543)	ne na kalej yele e saya kena karajeke e cikib yale sa saji
Gln422-Tyr435	(1537)	elegatel o para elegatelación caryon conflétely de selvir est
Arg426-Lys432	(1561)	dejer tejakulést ki élembat veletekke a taga a leti ki tejek fejekke).
Arg426-Gly431B	(1561)	$e(e,k_0,\chi_0 e m_0\chi_0,\chi_1,\chi_0,m_0\chi_0,\chi_1,\chi_1,\kappa_1,\kappa_2,\kappa_2,\kappa_3) = (\chi_0 e m_0\chi_0,\chi_0,\kappa_2,\chi_1,\kappa_2,\kappa_3) \frac{1}{2}$
Asn425-Lys432	(1555)	୕ୄ୵୷୕୷୷ଢ଼୶୳ଌଽୣ୵ଢ଼୕୶ଢ଼୕୳ୠଽ୵୶ୄ୕୶ଽ୕୕ୄ୕ଢ଼୷ୄ୕ଢ଼ଢ଼୕ଢ଼ଢ଼ଢ଼ଢ଼ୠଢ଼ୠୣ ୵୷ ୶୳ ୡ୕୷ ୡ୕ୣଌୡ
Consensus	(1561)	CGCAGCCTGACCCTGACCGTGCAGCCCGCCAGCTGCTGA
		1601 1640
Ile424-Ala433	(1589)	espendes:
Trp427-Gly431	(1601)	$e_{i,j}(c,c_{i+j,j},\gamma_{i+j,j+1},\ldots,c_{i+j,j+1},c_{i+j+1},c_{i+j,j+1},\ldots,c_{i+j,j+1},\ldots,c_{i+j,j+1},c_{i+j+1},\ldots,c_{i+j,j+1},\ldots,c_{i+j+1$
Gln422-Tyr435B	(1577)	en tit en vilge heigt, eighterkenten en harbyeige gebied bij
Arg426-Gly431	(1601)	control will eligible of the delegation of the control of the cont
Ile423-Met434	(1583)	ejele ejet i swerke erke erke erketskriv i vir vir ke ike evre evrikelele i il
Gln422-Tyr435	(1577)	ପିର୍ବିତ ନୂଷ ଓ ଅଧିକୃତ୍ୟ କଥା କଥା କଥା କଥା । ପ୍ରସ୍ତ୍ରୀତ ଓ ଓ ପ୍ରତ୍ତିକ ଓ ଅଧିକୃତ୍ୟ କଥା ।
Arg426-Lys432	(1601)	$c_{N}(r q _{Q}, h_{r}(r)_{Q}, r _{Q}, h_{r}(r)_{Q}, h_{r$
Arg426-Gly431B	(1601)	ૡ૽ૡૺૡૺૡૺ૽૽૱ૹૡૡૺૡ૱ૡૺૺઌ૱ૡ૽૽ૹ૱ઌ૽૽૽ૹ૱ઌ૽૽ઌ૱ઌ૽૽ઌ૽૽ઌ૽૽ઌ૽૽ઌ૽૽ૡ૽ૡ૽ૡ૽ૢૡૺૡૢ૽ૡ૽૽ૡૺૡ૽૱ઌ
Asn425-Lys432	(1595)	લુંકાલું લોકો, ૧૨૦ લેવડા વરા પર્લાકો, પર્લાવા, પર્લાવા, પર્લાવા, પ્રાપ્તાના કર્યો, લાંધા કરી છે. કર્યો છે. કર્યો
Consensus	(1601)	GCGGCATCGTGCAGCAGCAGAACAACCTGCTGCGCGCCAT
	•	1641 1680
Ile424-Ala433	(1629)	શક્યાલા પ્રાથમિક ક્ષેત્ર કાર્યાલા કાર્યાલા કાર્યાલા કાર્યાલા કાર્યાલા કાર્યાલા કાર્યાલા કાર્યાલા કાર્યાલા કાર્ય
Trp427-Gly431	(1641)	ert.veres xxxyelxyycix seleniyelxiyelxiyelxiyelxienyelxiqeyideyide
Gln422-Tyr435B	(1617)	ૹૺૡ૽ઌ૾ૡૺૡ૽૽ૹ૽ૹૹઌૡ૱ઌઌઌઌઌઌ૽ૡ૽ૡ૽ઌ૽ઌૡઌૹઌ૽૽ૺઌૹૡૡ૽ૺઌ૾ૡૺૡ૽ૺ૱ૹૡ૽૽ૡ૽ૺૡૺૡૺ૱ૹૡ૽૽ૡ૽૽ૡ૽ૺૡૺ
Arg426-Gly431	(1641)	ल्ड म्युद्धिल्ल्र्ड्डम्यूल्यक्तम् वर्णम् । १८०० महावर्णम् । १८०० महावर्णम् ।
_		FIC ALL

FIG. 4H

		23 / 03
Ile423-Met434	(1623)	ેલપ્પાદાન કોર્ગ પ્પદારમ ાનું અમેલ ઉપયોગના મહિલા મહી આ કે મહિલા છે. મહામાનું દુવના લ
Gln422-Tyr435	(1617)	micstraction earlier articles the which the cortes entire the excellent entire
Arg426-Lys432	(1641)	स्टर्भरहास्त्रहोत्। स्थर (च्यरास्टर्भर (च्यरास्टर्भर (द्यर्गस्टर्भस्टर्भस्टर्भस्टर्भस्टर्भस्टर्भस्टर्भस्टर्भस्
Arg426-Gly431B	(1641)	s(c). Activities and control of the
Asn425-Lys432	(1635)	racentines experiently in established enterestrates from established enterestrates and establish
Consensus	(1641)	CGAGGCCCAGCAGCAGCTGCAGCTGACCGTGTGGGGC
		1681 1720
Ile424-Ala433	(1669)	where very desident delegated and outside a soft electrical and
Trp427-Gly431	(1681)	takentatur entetem kenenturjena erelejejean elem elem elem elem elem elem elem ele
Gln422-Tyr435B	(1657)	: अग्रेर-१५५४दाल-५८दाल-१५दालकः अद्युक्तः । व्यवस्थितः । व
Arg426-Gly431	(1681)	thicks cover a market delactates of the content of
Ile423-Met434	(1663)	୵୰୰୷୳୰୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷
Gln422-Tyr435	(1657)	ନ୍ଧାନୀୟ (୧୯୯୯) କର୍ଷ୍ଟ ଜାପ୍ତର୍ଶ୍ୱର୍ଷ୍ଟ ହେବ ହେବାର୍ଷ୍ଟ ଖିତ (ପ୍ରି ଖ ୀୟ) ହେବା
Arg426-Lys432	(1681)	Angles in yell in the located of the exception of the exception of the extension of the ext
Arg426-Gly431B	(1681)	At a taticies y clessicioniticiciologici de la celentrici de esta dicital de cara
Asn425-Lys432	(1675)	Asset is a Constitution of the second of the second of the control of the second of th
Consensus	(1681)	ATCAAGCAGCTGCAGGCCCGTGCTGGCCGTGGAGCGCT
T1 404 T1 400		1721 1760
Ile424-Ala433	(1709)	97. wipkiejs/ejs/kon/kreiphips/kop/kop/kejs/kejs/kejs/kop/kejs/kop/kejs/kejs/kejs/kejs/kejs/kejs/kejs/kejs
Trp427-Gly431	(1721)	$\overline{\mathcal{A}_{G}}: \exists_{i} c_{i+1}, x_{i}/C_{i} c_{i+1}/C_{i}, c_{i}/C_{i} \otimes x_{i}/C_{i} \otimes \exists_{i} c_{i} c_{i} + c_{i}/C_{i} \otimes x_{i}/C_{i} \otimes x_{i}/C_{$
Gln422-Tyr435B	(1697)	facetaretarretes (electric machinetes) electric machinetes (electric mac
Arg426-Gly431	(1721)	voca, svyvelove syvelovoje neje prejedevnije skoje nejedevnije ve
Ile423-Met434	(1703)	र्वक्षाम् १८८५/५५१ हः । १८० १८७म् । इ.स. १५५ <mark>५ १५७ १५५ १५५५ । १५५५५ १५५५ १५५५</mark>
Gln422-Tyr435	(1697)	୍ରତ୍ତ (୧୯୯) ମଧ୍ୟ ପ୍ରକ୍ରେ ଅବସ୍ଥାର ଅନ୍ୟୁକ୍ତ ହେଉଛି । ଅନ୍ୟୁକ୍ତ ହେଉଛି
Arg426-Lys432	(1721)	२, वरण पटः भगविद्यः २, ४, १, १८ वर्षः भविष्यः १, १८ वर्षः १, १८ वर्षः १, १८ वर्षः १, १८ वर्षः १, १८ वर्षः १, १८
Arg426-Gly431B	(1721)	Acominante caronación de reigile de la refleie de mileo. E
Asn425-Lys432	(1715)	Aconica Account to Account to Section (Account to Account to Accou
Consensus	(1721)	ACCTGAAGGACCAGCTGCTGGGCATCTGGGGCTGCAG
Ile424-Ala433	(1740)	1761 1800
Trp427-Gly431	(1749) (1761)	DESCRIPTION OF THE PROPERTY OF
Gln422-Tyr435B	(1737)	e[e]e[e]e]e e e e e e e e e e e e e e e
Arg426-Gly431	(1761)	ere a ser a ser a ser l'attaine de
Ile423-Met434	(1743)	AGLERATE INC. II. REST STATE COLOR SE LEGISTATE :
Gln422-Tyr435	(1737)	and the second second and the second of the
Arg426-Lys432	(1761)	garanta ta sa
Arg426-Gly431B	(1761)	સંવેદાદ પ્રાપ્ત માન્ય ભાગમાં હવા જ મુંદ્રોકા વાર્ષ આ મેલિએમ્પ્રે પ્રદા ર્શ
Asn425-Lys432	(1755)	हें शहर : (एहं: :६ म हे (हहा) () है । हो है (एहं: सह्दर्भ में () में ()
Consensus	(1761)	CGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCC
Compensus	(1701)	1801 1840
Ile424-Ala433	(1789)	'(Contectivities and additional Contestion of Contestion o
Trp427-Gly431	(1801)	Actor of the respondent securities of the contract of the cont
Gln422-Tyr435B	(1777)	Nefabro Ano an arthuran of a coept to the competition of
Arg426-Gly431	(1801)	\$\text{\tex{\tex
Ile423-Met434	(1783)	
Gln422-Tyr435	(1777)	Agt and Veloviasi taina anglan sanasassigi anglanat Agt and Veloviasi taina anglan sanasassigi anglanat
Arg426-Lys432	(1801)	taje kajastajaja, tostatlasijajan kalastajastsatalisikajastijastsiost
Arg426-Gly431B	(1801)	.(c:::///cit/:/(c/c);/:://cit/://cit/://cit//cit/://cit////cit//cit
Asn425-Lys432	(1795)	\$\c.\c.\c\c\c\c\c\c\c\c\c\c\c\c\c\c\c\c\
Consensus	(1801)	AGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAACA
		1841 1880
Ile424-Ala433	(1829)	े स्वरं र कर हाराम् र ता र स्वरं हात्र स्वतंत्र स्वरं स्वरं स्वरं कर स्वरं कर स्वरं कर स्वरं कर स्वरं कर स्वरं
Trp427-Gly431	(1841)	ntert, o contrator e les transleteres trefetetet en en recenteira minimo en de
Gln422-Tyr435B	(1817)	ketytoen elesin elesiden eleletyteleseleletytelistytylljelesidosainmikytes (e
Arg426-Gly431	(1841)	Kedtyon v elektyne en rein kintertytelolen en reinning och (varstion) kintertyterik
Ile423-Met434	(1823)	ૡૡઌૡૡૡૡૡૡૡૡૡૡૡૡૡૡૡૡૡૡૡૡૡૡૡૡૡૡૡૡૡૡૡૡૡૡ
Gln422-Tyr435	(1817)	Retriefestigles et este este est anne et est est est (test anne en en este est (test est est est est est est e
		FIO AF

FIG. 4I

Arg426-Lys432	(1841)	wertelenkelentakeleitkenkelelettelens entrettakelettentation
Arg426-Gly431B	(1841)	inglitaro, ingles, angla sepinaletes selestes es statos y a les ves a variabiles se
Asn425-Lys432	(1835)	Retaction (elegant description and the second control of the secon
Consensus	(1841)	TGACCTGGATGGAGTGGGAGCGCGAGATCGACAACTACAC
		1881 1920
Ile424-Ala433	(1869)	અજ્ઞાતાઓ મુક્તા, અહું અલ્લા, કર્યા છે. જે અલ્લા, અનું અલ્લા, વ્યાવના, વ્યાવના, વ્યાવના, અલ્લા, જે પછા છે. (હ
Trp427-Gly431	(1881)	# : ACCURC # : JPA'01,1000 # : AVEOCRACION/CHACO # : MARCONAC
Gln422-Tyr435B	(1857)	et. V. (v.) et. et. et. et. et. (c.) etek et., et. (et.) et., (et.) et., (et.) et., (et.) et., (et.) et., (et.)
Arg426-Gly431	(1881)	e1;v;v:v:e=vex:\v:v:v:xy(e=v;v=v=v=v=v=v=v=v=xy=y=v=v=xy=xy=y=v=v=xy=y=v=v=xy=y=v=v=xy=y=v=v=xy=y=v=v=xy=y=xy=v=xy=xy=xy=xy=xy=xy=xy=xy=xy=xy=xy=xy=xy=
Ile423-Met434	(1863)	er saktologiden, kisikkopikos, koloko ilikeritukolenikeise kierikete osa kirila, kolokike
Gln422-Tyr435	(1857)	stratele Keathyleantaich sin sin eile antaite istat eile achtaire an Athai
Arg426-Lys432	(1881)	with the largest transfer with the m is estimated the content of the M in the M
Arg426-Gly431B	(1881)	and telephone and extraction of contents followed telephone to the trace of the property of the trace of the
Asn425-Lys432	(1875)	୭୩/୬୯୧: ଏହି ଏହି ଅନ୍ତର୍ଶ ବର୍ଷ ବର୍ଷ ପ୍ରଥମ କରିଥିବା ହେଉଥିଲି ।
Consensus	(1881)	CAACCTGATCTACACCCTGATCGAGGAGGCCAGAACCAG
		1921 1960
Ile424-Ala433	(1909)	ercic (etratorial), antelercia realizative alle allegative de la missoria
Trp427-Gly431	(1921)	$w_i(c;c_i)(c_i^{i})_{i=1}(c), x_i(c;c_i)(c c_i)(c c_i)(c c_i)(c) = (c c_i)(c c_i)(c) = (c c_i)(c c$
Gln422-Tyr435B	(1897)	ૡૡૺૡૡઌ૽૽૱ઌઌ૽૽૱ૡૡ૽ૡૡ૱૱૱ૡૡૡૡ૱૱૱ <mark>ૡઌ</mark> ૽ૢ૽ૡ૽ૢ
Arg426-Gly431	(1921)	લા (૮)લા, (લા) પ્રાપ્ત અપૂર્ણ હોય છે. જે કે મુંદ્રા હોય છે. માને લાગ કરે કરા માને લાગ કરે છે. માને લાગ કરે કે
Ile423-Met434	(1903)	stycles/tesphalphycles/tesphere/februse/lebrusels/tesphere/februse/februse/
Gln422-Tyr435	(1897)	Michigan and Control of the Control
Arg426-Lys432	(1921)	wateratera in a tolera watera ee en
Arg426-Gly431B	(1921)	મતાલાતામાં ભાગો હતા. આવાલાતાના પ્રાથમ માના
Asn425-Lys432	(1915)	water cut/cut/styletys cut/clestric style entrestricts/yelestricts/yelein
Consensus	(1921)	CAGGAGAAGAACGAGCAGGAGCTGCAGAGCTGGACAAGT
		1961 2000
Ile424-Ala433	(1949)	ej melet, Velelekter (ejen), mitelette ett ett, million (k. m.) ett olij
Trp427-Gly431	(1961)	ଗ୍ରମ ନ୍ୟୁର ଅନ୍ତର ଅନ୍ତର କ୍ଷିତ୍ର ଅନ୍ତର ଅନ୍ତର ଅନ୍ତର । ପ୍ରତ୍ୟ ସମ୍ବର୍ଜ ଅନ୍ତର ଅନ୍ତର ଅନ୍ତର ଅନ୍ତର ଅନ୍ତର ହେଉଛି ।
Gln422-Tyr435B	(1937)	ccc.cc acc atematers, and a cac
Arg426-Gly431	(1961)	लेड इंडिक्सर्वाहोडी इस्तिक अस्ति १००० हर्ने १००५ वर्ग्सर १५०० हर्ने १५०० हर्ने १५०० हर्ने १५० हर्ने १५० हर्ने
Ile423-Met434	(1943)	ejejustratoris ir iejenato televių ir ori aktivateria karalėje ji
Gln422-Tyr435	(1937)	Table de la company de la comp
Arg426-Lys432	(1961)	લાનું નું લાગામાં કરવામાં માટે કરાય કાર્યા છે. તે
Arg426-Gly431B	(1961)	वद्वतः । तरसः । १० । १७ । ११ । ११ वः । १९ । ११ । १४ । १४ । १४ । १४ । १४ । १४
Asn425-Lys432 Consensus	(1955)	ANCE CANCEL FOR THE PARKS OF THE CONTROL OF THE PARKS OF THE PARKS OF THE CONTROL OF THE PARKS O
Consensus	(1961)	GGGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCT 2001 2040
Ile424-Ala433	(1989)	2001 2040
Trp427-Gly431	(2001)	
Gln422-Tyr435B	(1977)	EMPCIC BIVE STATES AND STATES OF SERVICE SERVICES OF S
Arg426-Gly431	(2001)	ringlianto thought a confort of the state of sield (sield is
Ile423-Met434	(1983)	ch chartentariates a might carried in the called block before by
Gln422-Tyr435	(1977)	epitefoli (2) (etc.) - vyroza ytelik i Gyyddyn die i i i i i er efejt efek i by <mark>e</mark>
Arg426-Lys432	(2001)	ent in the service of
Arg426-Gly431B	(2001)	entelshar(entaratish) amiliketare@j.astros etcl.icidetel. s
Asn425-Lys432	(1995)	Chrescia (astrografica com marso e digiral com cipe e elejeteje de
Consensus	(2001)	GTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTG
	(2001)	2041 2080
Ile424-Ala433	(2029)	arcidedericaderraetricieraterant-crattericieratere
Trp427-Gly431	(2041)	epiterelation decembritace depiter a delectric contentation and expedit
Gln422-Tyr435B	(2017)	છા કેલ (ફોલ્ટ) કેલ કેલ લેવા કરો એ કોલ કેલ કેલ કેલ કેલ કેલ કેલ કેલ કેલ કેલ કે
Arg426-Gly431	(2041)	chrelete in the old or debression of the selective letter of the est
Ile423-Met434	(2023)	eyenicicisi nicinicisi. en mehrofimostatorienroforestatorien del model de f
Gln422-Tyr435	(2017)	en neg trene conference a ponte for suffering to the conference of the suffering of the suf
Arg426-Lys432	(2041)	epidojojojoja sidelo ojoj, odoje pojembiljen Vermentolejito na tom integraljen
Arg426-Gly431B	(2041)	distributed for the content of the c
=	•	

FIG. 4J

Asn425-Lys432	(2035)	chreleicheimemerer mer breihmervollet beibberveichtbie beib
Consensus	(2041)	GTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGA
		2081 2120
Ile424-Ala433	(2069)	र,लंलब्बदमहत्वबद्धल्केर्द्रद्विद्धलार्भक्षेत्रकार्द्धार हार्द्धार हार्द्धार स्थापन
Trp427-Gly431	(2081)	Action chile exists to element to the contract of the major the contract of th
Gln422-Tyr435B	(2057)	કુલુલુલુલુલ્યુલ્યલું સાથે તાલુકા છે. કુલુલુલું કુલુલું કુલુલું કુલુલું કુલ કુલુલું કુલ કુલુલું કુલ કુલુલું કુલ
Arg426-Gly431	(2081)	ୀ,ଜାନାପ୍ରମୟମୟର ପାର୍ଟ୍ୟ,(ଜାମ୍ୟାପ୍ରମୟ,୧୯୬) ଓ ଅଟନ୍ତ ବ୍ୟୟକ୍ତମ୍ୟନ୍ତ (କ୍ୟମନ୍ତ (କ୍ୟମନ୍ତ (କ୍ୟମନ୍ତ (କ୍ୟମନ୍ତ (କ୍ୟମନ୍ତ (କ
Ile423-Met434	(2063)	भूतिवाद्यत्तरत्त्वात्त्रवात्रवात्त्रवात्त्रवात्त्रवात्त्रवात्त्रवात्त्रवात्त्रवात्त्रवात्त्रवात्त्रवात्त्रवात्
Gln422-Tyr435	(2057)	wedges contientendenten, streterentententententen
Arg426-Lys432	(2081)	16, Louis Control of C
Arg426-Gly431B	(2081)	More larger that region is to let a mixter it also reconfigurate which receives
Asn425-Lys432	(2075)	ॱ१७:बद्रबद्धनं द्वर,संबंबक्षप्रसद्धिक ११४३ ।
Consensus	(2081)	ACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCCAGAC
-1 404 -1 4-4		2121 2160
Ile424-Ala433	(2109)	ଶ୍ୟରତୀ । ପ୍ରତିଷ୍ଠିତ୍ୟରତ ୍ୱ ଅବସ୍ଥର୍ଗର ପ୍ରତିଷ୍ଠ (ବ୍ୟ <mark>ର୍ଗର ଗ୍ରିମ୍ବର</mark> ଅ <mark>କ୍ୟାପ୍ର ।</mark>
Trp427-Gly431	(2121)	ପ୍ରପ୍ର : rc/ଟ୬୪/ପ୍ରମ୍ବିମ୍ବେଟ ୧୯୯୦ : ୧୯୯୭ ଅଟଲ:/ସ୍କ୍ୟୁମ୍ବେଟ ୧୯.୪୯ମସ୍କ
Gln422-Tyr435B	(2097)	ଖଣ୍ଟ ଧଳ ଜଣ ପସ୍ତ୍ରେ ଲୋକ ସେମ୍ବର ଅନ୍ତମୟ ପ୍ରତ୍ରେ ମସର ପ୍ରଥମ ।
Arg426-Gly431	(2121)	ଶ୍ୟୁମ୍ନ : : ଜର୍ଗ ଓ ନର୍ଜ୍ଞାନ ଧାରଣ ବଳ ଅଧ୍ୟର୍ଗ <mark>ମ</mark> ଓ ଗ୍ରେମ୍ବର ଜଣ୍ଣ ଖଣ୍ଡ ନ୍
Ile423-Met434	(2103)	ब्रहर कोहर ४ व्हारास ४ व्हार समित का समित विवास विवास ।
Gln422-Tyr435	(2097)	ବ୍ୟ ପ୍ରତିଶ୍ୱର ପ୍ରତିଶ
Arg426-Lys432	(2121)	Marchen internal of the content of t
Arg426-Gly431B	(2121)	ବ୍ୟାତ୍ୟ : ଖଣ୍ଡପ୍ରସ୍ତ ପ୍ରମିଶ ଓଡ଼ିଆ ହୋଇଥିଲି ଅଟ୍ରିଗ୍ରିମ୍ୟ ଅଟନ୍
Asn425-Lys432	(2115)	୬ ଜଣିବା ଓ ପର୍ବ ଭା <mark>ଗି ଓଡ଼ିଆ ଓଡ଼ିଆ ଓଡ଼ିଆ ଜଣିକ ଜଣି । ଜଣିକ ଜଣି ଓଡ଼ିଆ ଜଣିକ ଜଣିକ ଜଣିକ ଜଣିକ ଜଣିକ ଜଣିକ ଜଣିକ ଜଣିକ</mark>
Consensus	(2121)	CCGCTTCCCCGCCCCCGCGCCCCGACGCCCCGAGGGC
T10424 N1-422	(01.40)	2161 2200
Ile424-Ala433	(2149)	, ଧୂର ୧୯.୬୧୯୮୬,୧୯୯୯,୧୯୯୯,୧୯୯୯,୧୯୯୯,୧୯୯୯,୭୯୯,୭୯୯,୭୯୯,୭୯
Trp427-Gly431	(2161)	Micci/cu/jet/dacentamicseranoeleit/eeleit/cel
Gln422-Tyr435B	(2137)	୬ ୬ ୬ ୬ ୬ ୯୯୬ ୬ ୯୪ ୯ ଓଡ଼େ ପ୍ରକ୍ରେମ୍ବର ପ୍ରକ୍ରିକ ଓଡ଼ି
Arg426-Gly431 Ile423-Met434	(2161)	$\label{eq:continuity} \mathcal{A}_{\mathcal{A}}(x) = \{ (x,y) \in \mathcal{A}_{\mathcal{A}}(x) : (x,$
Gln422-Tyr435	(2143)	<u>ම් මේට්ටේ වස්ට්ට්ටේ මේට්ට්ටේමට මේට්ට්ට්ට්ට්ට්ට්</u>
	(2137)	भारतम् । तस्यतिसर्वे सर्वे साम् विकासिकाः अस्तरः वास्तरः यह स्वरः स्वरः स्वरः
Arg426-Lys432 Arg426-Gly431B	(2161)	A Zelts wiese Gestelle zu der 1953 der 1955 Gestelle Gestelle Gestelle G
Asn425-Lys432	(2161)	. ଜ୍ଞାକଟ-ସମ୍ବାଦ ଅନ୍ତମ୍ଭ ପର୍ବ ସମୟ ହେଉଟ ଅନ୍ତମ୍ଭ କଥା ହିଲା ହିଲା ଅନ୍ତମ୍ମ ।
Consensus	(2155) (2161)	AMOCA CA CCA CCA CCA CCA CCA CCA CCA CCA C
Consensus	(2101)	ATCGAĞGAGGAGGGCGAĞCGCGACCGCAGCA 2201 2240
Ile424-Ala433	(2100)	
Trp427-Gly431	(2189) (2201)	ଗ୍ରିଗର ଅଟେ ଏକ୍ଟର ଅଟେ ବେଟେ ଏକ୍ଟର ଅଟେ ଏହା ହେଛି । ଅଧାର ପ୍ରତ୍ୟୁକ୍ତ ଅଟେ
Gln422-Tyr435B	(2177)	dele, esta del no accidante este a esta en accidante est esta esta esta esta esta esta est
Arg426-Gly431	(2201)	efection of the constant of the charter discovered state of the charter of
Ile423-Met434	(2183)	विश्वास्त्राम् । भूर दोष्टाके प्रदार हो १० १६ । । १८१६ । १ ६५५६) में १८ मुझेटीके प्र
Gln422-Tyr435	(2103)	tion of figure colors of the description of the property of the
Arg426-Lys432	(2201)	To the construction of the
Arg426-Gly431B	(2201)	
Asn425-Lys432	(2195)	compressing adology in the control of the control o
Consensus	(2201)	GCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACGA
	\2201)	2241 2280
Ile424-Ala433	(2229)	a (som the long that the contraction in the same and a little of the contraction in the same and a little of the contraction in the same and a little of the contraction in the same and a little of the contraction in the same and a little of the contraction in the same and a little of the contraction in the same and a little of the contraction in the same and a little of the contraction in the same and a little of the contraction in the same and a little of the contraction in the same and a little of the contraction in the same and a little of the contraction in the same and a little of the contraction in the same and a little of the contraction in the same and a little of the contraction in the same and a little of the contraction in the same and a little of the contraction in the same and a little of the contraction in the same and a little of the contraction in the contrac
Trp427-Gly431	(2241)	west cateful activities with the first activity of the contraction of
Gln422-Tyr435B	(2217)	national and the solitical contraction of the contr
Arg426-Gly431	(2241)	elenktivita in varionika kitaina kanka kata kata kata kata kata kata k
Ile423-Met434	(2223)	gichrigerantelechtentelmenntalen hitreine ales gilligeree
Gln422-Tyr435	(2217)	elenificas(elenificate) entraffato (entraformitato) (entraformitate) elenificate (entraformitation) entraffato
Arg426-Lys432	(2241)	ological selection of the selection of t
Arg426-Gly431B	(2241)	નું ભાષા લાક્ષ્ય કર્યા છે. જે
Asn425-Lys432	(2235)	cleanfelocationsenrametermenten en en en mile perfection elle
Consensus	(2241)	CCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCCTGCGC
	/	F10 414

FIG. 4K

		2281 2320
Ile424-Ala433	(2269)	त्र, त्रवारतः प्रतिवृद्धे रक्षः यः वृद्धिवृद्धाः वृद्धिः । प्रतिवृद्धाः । प्रतिवृद्धाः । प्रतिवृद्धाः । प्रतिव 2020
Trp427-Gly431	(2281)	જોડ(લોલોમોલી: પોલીલોપ્રેલોડ)પાલોલોલોલા લોક કોલીલોપાય કોલીલોપાલોલો પાલોલો
Gln422-Tyr435B	(2257)	et Ac chilet, microhilet i procede colet chile of the weather selection selections
Arg426-Gly431	(2281)	etweletwerkniegenwerkniegen og veletetervingehveletweletwele
Ile423-Met434	(2263)	atyoktyaty sienkawindeletateethie athroestalane siglamethiele
Gln422-Tyr435	(2257)	c),૧૯૮૦ પ્રદેશ, સ્મૃત અમૃ દ ્યાં પૂર્વ કર્યા હતા. હોલા છે. જે જે એ લેવો મૃત્યું લોકો પ્રદો લો પ્ર દો લો પ્રદોલો પ્ર
Arg426-Lys432	(2281)	જે.પુંચલ ૧૯,૫૫૦ હોમના પાંચલલ હાર લેવા હાલના પાંચલ ા પાંચલનો પાંચલનો પાંચ
Arg426-Gly431B	(2281)	canalante nationality of the therm of converte states and converted
Asn425-Lys432	(2275)	CLUSSINGLY INDVENTARISCONCINCIONO CONTARIO DE CONTARIO CONTARIO DE CONTRARIO DE CONTRARIO CONTRA
Consensus	(2281)	GACCTGATCCTGATCGCCGCCCCGCATCGTGGAGCTGCTGG
		2321 2360
Ile424-Ala433	(2309)	cle activities and a responsibility of the second of the s
Trp427-Gly431	(2321)	ଗ୍ରୀଣ୍ଟାର୍ଗ୍ରମ୍ମ ବ୍ୟୁମ୍ନ ମଧ୍ୟ କ୍ଷ୍ମ ବ୍ୟୁମ୍ନ କ୍ଷ୍ମ ମଧ୍ୟ କ୍ଷ୍ମ କ୍ଷ୍ମ କ୍ଷ୍ମ କ୍ଷ୍ମ କ୍ଷ୍ମ କ୍ଷ୍ମ କ୍ଷ୍ମ କ୍ଷ୍ମ କ୍ଷ୍ମ କ
Gln422-Tyr435B	(2297)	enoge compressive entrans of the first our energy and the first our ene
Arg426-Gly431	(2321)	chemiclopic control de la
Ile423-Met434	(2303)	eleleleleneleleleleneleleratetet elenterational/antereleratea/
Gln422-Tyr435	(2297)	द्विहाद्यकृत्यक्षः अस्य स्थापन्य स्थापन्य स्थापन्य स्थापन्य स्थापन्य स्थापन्य स्थापन्य स्थापन्य स्थापन्य स्थापन
Arg426-Lys432	(2321)	e matieur : en relete rete elle de silessités rélete elle ell
Arg426-Gly431B	(2321)	TRECONSEQUENT TECHNOLOGIC STRUCTOR CONFERENCE AND THE
Asn425-Lys432	(2315)	તિ લાકાર કાર્યકાલ (અપનાલાના (લાકાર કાર્યકાર, કાર્યકાર ક્રોકાર કાર્યકાર, કર્યકાર કર્યા છે. આ
Consensus	(2321)	GCCGCCGCGCTGGGGGCCCTGAAGTACTGGGGCAACCT
		2361 2400
Ile424-Ala433	(2349)	୯୩୮ ଅନ୍ୟୁକ୍ତ ହେଉଁ ବର୍ଷ ଅନ୍ୟୁକ୍ତ ହେଉଁ ହେଉଁ ହେଉଁ
Trp427-Gly431	(2361)	ero, en stancióa e degeneros anos en el superiorizado de exigencia
Gln422-Tyr435B	(2337)	FORETYCE CONTROL OF THE CONTROL
Arg426-Gly431	(2361)	en velstigenten men hijn erteretien der mer viroliken med de re
Ile423-Met434	(2343)	CONTRACTOR OF CO
Gln422-Tyr435	(2337)	sjelijejera i za stelerav opraejeracja ujelizacijanoga sleje duje
Arg426-Lys432	(2361)	CGF_CC En VIII 61, CEF_EFL EXYENVENTEE EL E
Arg426-Gly431B	(2361)	e so edos Ache, osuedo sulo ed Colo. Te so les Mentes Acho, e e
Asn425-Lys432	(2355)	cia a color de dele de destación del alces a feliplas del elegio del co
Consensus	(2361)	GCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTG
		2401 2440
Ile424-Ala433	(2389)	te covern enteles procedes and confederations of
Trp427-Gly431	(2401)	necessie en de la motor un legiode concominete à
Gln422-Tyr435B	(2377)	ମନ୍ତ୍ର ଓ ୧୯୯୯ ମନ୍ତି ମନ୍ତିଆନ ମନ୍ତି ହେ ହେ <mark>ନ୍ତିର ଅନ୍ତର ଅନ୍ତିଆନ</mark>
Arg426-Gly431	(2401)	se e fet stefes et de Antel e de la compete et de
Ile423-Met434	(2383)	ret i del rete estade di mancio della con della estade estade estade e della estade e della estade e della est
Gln422-Tyr435	(2377)	preferance of rest to the electric designer of the second
Arg426-Lys432	(2401)	ଏମିକ ଆଧାର ଓ ଜଣ ପ୍ରମହିତ । ମ ପ୍ରମହିତ ହେଉଛି ଅନ୍ୟର୍ଶ ହେଉଛି
Arg426-Gly431B	(2401)	to de la Weletto de ville ele
Asn425-Lys432	(2395)	Act :
Consensus	(2401)	AGCCTGTTCGACGCCATCGCCATCGCCGTGGCCGAGGGCA
		2441 2480
Ile424-Ala433	(2429)	୵୶୶୵୕୵୶ୡ୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷
Trp427-Gly431	(2441)	ଝିରାଫ (ଆବ୍ରମ୍ୟ ମଧ୍ୟର ଓଡ଼ି (ଜାନ ୍ମ ମହର ୪ ଲୀ ମହିଳ୍କ (ବ୍ୟରଣ ପ୍ରିଟ ଅଟର ଓଡ଼ି
Gln422-Tyr435B	(2417)	ଧିକ :>\\ାଧାର୍ଯ୍ୟ ମୁଖ୍ୟ ଜ୍ୟେ : ୧୯୧୯ (ଜ୍ୟୁ ମୁଖ୍ୟ ଜ୍ୟୁ ମୁଖ୍ୟ ଜ୍ୟୁ ମୁଖ୍ୟ ଜ୍ୟୁ ମୁଖ୍ୟ ଜ୍ୟୁ ମୁଖ୍ୟ ଜ୍ୟୁ ମୁଖ୍ୟ ଜ୍ୟୁ ମୁଖ
Arg426-Gly431	(2441)	୬୬.୧୯ (୬୮୯୮) ୬୬.୧୯/୪୭.୧୯/୪୯୮୬ : ମସ୍ତର୍ଗ ୬.୧୯/୪୯ ୦ (୬୯୯୮) ଅଟି
Ile423-Met434	(2423)	ୣ୷ୠ୷୕୰ୡ୕ୣ୷୷୷୷୷୷ୡ୕୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷
Gln422-Tyr435	(2417)	ସ୍ତ୍ରକ୍ତ (୧୯୯ ପ୍ରକ୍ୟ (୧୯) । ଅନ୍ୟୁକ୍ତ ଅନ୍ୟୁକ୍ତ ଅନ୍ୟୁକ୍ତ (୧୯) ଅନ୍ୟୁକ୍ତ ଅନ୍ୟୁକ୍ତ ଅନ୍ୟୁକ୍ତ ।
Arg426-Lys432	(2441)	विविद्यः, श्रेष्टाः, वर्षेक्षः, १००१, १०१८ वर्षे, १००० वर्षे, १००० वर्षे ।
Arg426-Gly431B	(2441)	द्रश्चारतित्वात्रात्ताक्षरातः स्तरतितः स्तर्वहत्त्वः स्तिविद्याः स्तरितः । तरस्र
Asn425-Lys432	(2435)	୷୲୕ଌ୵୲୶୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷
Consensus	(2441)	CCGACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCGC
		2481 2520
Ile424-Ala433	(2469)	changescharzen arrientelaren a. eta teralet hafet eta teralet elegen kelebre
		FIG AI

FIG. 4L

Trp427-Gly431	(2481)	ામાનક અમ ન ભ^{ર્}કા . કહેલા કહેલા છે. અંદાવિસ્તાન સંવેદ સંપ્રાને કોર્યા છે. છે. જે માને સ્ત્રામાં સ્ત્રામાં માને સ્ત્રામાં સ્ત્રામાં માને સ્ત્રામાં માને સ્ત્રામાં માને સ્ત્રામાં સ્ત્રામા સ્ત્રામા સ્ત્રામા સ્ત્રામાં સ્ત્રામા સ્ત્રામાં સ્ત્રામા સ્ત્રામા સ્ત્રામા સ્ત્ર
Gln422-Tyr435B	(2457)	enthespiretisternices cedescentineses intereste
Arg426-Gly431	(2481)	An actual curve where the electricity of the fide in the cive
Ile423-Met434	(2463)	ઌ૾ૻૹ૾ૹૡઌ૽ ૽ઌ૽ ૱ૹઌઌ૽૽ૼ૽૽૽૽૽ઌૡ૽૽૽ૢઌ૽ૻઌ૽૽ૹ૽ૹ <mark>ઌ૽૽૽ૢૡૡૡ</mark> ઌ૽ૹૣ૽ૢૢૢૢ૽ૢ૽ૢઌ૽૽ૺ૾ૢૺૡ
Gln422-Tyr435	(2457)	saping to intelest the rate of electric connections the billion of the
Arg426-Lys432	(2481)	PARTICIONALESTA O ANNA O CONTRACESTANTA DE DOS PERSONALES (CONTRACES
Arg426-Gly431B	(2481)	PRINCIPLE ON CONTROL OF CONTROL O
Asn425-Lys432	(2475)	elikikolehkeletivatyisele elikieleteletenikieleteletenikieletele
Consensus	(2481)	CTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAG -
		2521 2541
Ile424-Ala433	(2509)	ejerrieteide jade terrateta et provincio existeri
Trp427-Gly431	(2521)	nantagentagenten size och E
Gln422-Tyr435B	(2497)	of conference, he mindentally them on the
Arg426-Gly431	(2521)	MERCECOCCUTOR SCREAM STEELE
Ile423-Met434	(2503)	ofelectorely, here wiles by your extensive to the
Gln422-Tyr435	(2497)	NEWC NO CHIEN REPORTATION ROLE IN
Arg426-Lys432	(2521)	Grief: Gradeling and Alberta Are
Arg426-Gly431B	(2521)	CGCGCCGTECTEGTAATTICCAG
Asn425-Lys432	(2515)	PERENGERICE CLEVY.
Consensus	(2521)	CGCGCCCTGCTGTAACTCGAG

FIG. 4M

WO 00/39303	28	/ ₁ 65 PCT/US99/31272
		1 30
Leu122-Ser199-Tryp427-Gly431		GAATTCGCCACCATGGATGCAATGAAGAGA
Val127-Asn195-Arg426-Gly431		GAATTCGCCACCATGGATGCAATGAAGAGA
Val120-Thr202-Ile424-Ala433	(1)	The second section of the sect
Leul22-Ser199-Arg426-Lys432	(1)	GAATTCGCCACCATGGATGCAATGAAGAGA
Leu122-Ser199-Arg426-Gly431	(1)	GAATTCGCCACCATGGATGCAATGAAGAGA
Lys121-Va1200-Asn425-Lys432	(1)	(1) (1) (1) (1) (1) (1) (1) (1) (1) (1)
Val120-Ile201-Ile424-Ala433	(1)	GAATTCGCCACCATGGATGCAATGAAGAGA
Val120-Ile201B-Ile424-Ala433	(1)	GAATTCGCCACCATGGATGCAATGAAGAGA
Consensus	(1)	GAATTCGCCACCATGGATGCAATGAAGAGA
		31 60
Leu122-Ser199-Tryp427-Gly431	(31)	GGGCTCTGCTGTGTGCTGCTGTGTGGA
Val127-Asn195-Arg426-Gly431	(31)	GGGCTCTGCTGTGTGCTGCTGTGTGGA
Val120-Thr202-Ile424-Ala433	(31)	GGGCTCTGCTGTGTGCTGCTGTGTGGA
Leu122-Ser199-Arg426-Lys432	(31)	GGGCTCTGCTGTGTGCTGCTGTGTGGA
Leu122-Ser199-Arg426-Gly431	(31)	GGGCTCTGCTGTGTGCTGCTGTGTGGA
Lys121-Val200-Asn425-Lys432	(31)	GGGCTCTGCTGTGTGCTGCTGTGTGGA
Val120-Ile201-Ile424-Ala433	(31)	GGGCTCTGCTGTGTGCTGCTGTGTGGA
Val120-Ile201B-Ile424-Ala433	(31)	- Theodopeenhausting (1918-1918) - 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Consensus	(31)	GGGCTCTGCTGTGTGCTGCTGTGTGGA
Consensus	(31)	61 90
Leu122-Ser199-Tryp427-Gly431	(61)	GCAGTCTTCGTTTCGCCCAGCGCCGTGGAG
Val127-Asn195-Arg426-Gly431	(61)	GCAGTCTTCGTTTCGCCCAGCGCCGTGGAG
Val120-Thr202-Ile424-Ala433	(61)	The manufacture of the control of th
		GCAGTCTTCGTTTCGCCCAGCGCCGTGGAG
Leu122-Ser199-Arg426-Lys432	(61)	GCAGTCTTCGTTTCGCCCAGCGCCGTGGAG
Leu122-Ser199-Arg426-Gly431	(61)	GCAGTCTTCGCCCAGCGCCGTGGAG
Lys121-Val200-Asn425-Lys432	(61)	GCAGTCTTCGTTTCGCCCAGCGCCGTGGAG
Val120-Ile201-Ile424-Ala433	(61)	GCAGTCTTCGTTTCGCCCAGCGCCGTGGAG
Val120-Ile201B-Ile424-Ala433	(61)	GCAGTCTTCGTTTCGCCCAGCGCCGTGGAG
Consensus	(61)	GCAGTCTTCGTTTCGCCCAGCGCCGTGGAG
		91 120
Leu122-Ser199-Tryp427-Gly431	(91)	AAGCTGTGGGTGACCGTGTACTACGGCGTG
Val127-Asn195-Arg426-Gly431	(91)	AAGCTGTGGGTGACCGTGTACTACGGCGTG
Val120-Thr202-Ile424-Ala433	(91)	AAGCTGTGCGTGACCGTGTACTACGGCGTG
Leu122-Ser199-Arg426-Lys432	(91)	AAGCTGTGGGTGACCGTGTACTACGGCGTG
Leu122-Ser199-Arg426-Gly431	(91)	AAGCTGTGGGTGACCGTGTACTACGGCGTG
Lys121-Va1200-Asn425-Lys432	(91)	AAGCTGTGGGTGACCGTGTACTACGGCGTG
Val120-Ile201-Ile424-Ala433	(91)	AAGCTGTGGGTGACCGTGTACTACGGCGTG
Val120-Ile201B-Ile424-Ala433	(91)	AAGCTGTGGGTGACCGTGTACTACGGCGTG
Consensus	(91)	AAGCTGTGGGTGACCGTGTACTACGGCGTG
		121 150
Leu122-Ser199-Tryp427-Gly431	(121)	CCCGTGTGGAAGGAGGCCACCACCACCGTG
Val127-Asn195-Arg426-Gly431	(121)	CCGTGTGGAAGGAGGCCACCACCACCGTG
Val120-Thr202-Ile424-Ala433	(121)	CCCGTGTGGAAGGAGGCCACCACCACCCTG
Leu122-Ser199-Arg426-Lys432	(121)	CCCGTGTGGGAAGGAGGCCACCACCACCCTG
Leu122-Ser199-Arg426-Gly431	(121)	CCCGTGTGGAAGGAGGCCACCACCACCCTG
Lys121-Val200-Asn425-Lys432	(121)	CCCGTCTGGAAGGAGGCCACCACCACCCTG
Val120-Ile201-Ile424-Ala433	(121)	CCCGTGTGGAAGGAGGCCACCACCCCTG
Val120-Ile201B-Ile424-Ala433	(121)	CCCTTTGGAAGGAGGCCACCACCACCTTG
Consensus	(121)	CCCGTGTGGAAGGAGGCCACCACCACCCTG
Consensus	(444)	151 180
Leu122-Ser199-Tryp427-Gly431	(151)	TTCTGCGCCAGCGACGCCTACGAC
Val127-Asn195-Arg426-Gly431		TTCTGCGCCAGCGACGCCAAGGCCTACGAC
Vall20-Thr202-Ile424-Ala433	(151) (151)	to considerate and the termination of the second section of the second section of the second section is
		TTCTGCGCCAGCGACGCCAAGGCCTAGGAC
Leu122-Ser199-Arg426-Lys432	(151)	TTCTGCGCCAGCGACGCCAAGGCCTACGAC
Leu122-Ser199-Arg426-Gly431	(151)	TTCTGCGCCAGCGACGCCAAGGCCTACGAC
Lys121-Va1200-Asn425-Lys432	(151)	TTCTGCGCCAGCGACGCCAAGGCCTACGAC

WO 00/39303	29	/ 65	PCT/US99/31272
Val120-Ile201-Ile424-Ala433	(151)	TTCTGCGCCAGCGAC	GCCAAGGCCTACGAC
Val120-Ile201B-Ile424-Ala433	(151)	TTCTGCGCCAGCGAC	
Consensus	(151)	TTCTGCGCCAGCGAC	
oon bandab	(101)	181	210
Leu122-Ser199-Tryp427-Gly431	(181)	ACCGAGGTGCACAAC	
Vall27-Asn195-Arg426-Gly431	(181)	ACCGAGGTGCACAAC	
Val120-Thr202-Ile424-Ala433	(181)	Object Control (1) (4) 47 40 40 45 5 7 Min 2 March 100 40 40 40 40 40 40 40 40 40 40 40 40 4	ti i metriti i mi metriti i metrovi metrovina prekara prekara i transferi i transferi i m
Leu122-Ser199-Arg426-Lys432	(181)	ACCGAGGTGCACAAC	reproductive and productive and a reproductive productive producti
Leu122-Ser199-Arg426-Gly431	(181)	ACCGAGGTGCACAAC	And the second s
Lys121-Val200-Asn425-Lys432	(181)	ACCGAGGTGCACAAC	The state of the second st
Val120-Ile201-Ile424-Ala433	(181)	ACCGAGGTGCACAAC	The state of the s
Val120-Ile201B-Ile424-Ala433	(181)	ACCGAGGTGCACAAC	
Consensus		ACCGAGGTGCACAAC	
Consensus	(101)	211	240
Leu122-Ser199-Tryp427-Gly431	(211)	GCCTGCGTGCCCACC	=
Val127-Asn195-Arg426-Gly431	(211)	and the first throughout the authority of the second	4 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Val120-Thr202-Ile424-Ala433		GCCTGCGTGCCCACC	The state of the s
	(211)	GCCTGCGTGCCCACC	and the control of a control of a control of the co
Leu122-Ser199-Arg426-Lys432	(211)	GCCTGCGTGCCCACC	and the second s
Leu122-Ser199-Arg426-Gly431	(211)	GCCTGCGTGCCCACC	The state of the s
Lys121-Val200-Asn425-Lys432	(211)	A STATE OF THE STA	
Val120-Ile201-Ile424-Ala433	(211)	GCCTGCGTGCCCACC	
Vall20-Ile201B-Ile424-Ala433	(211)	GCCTGCGTGCCCACC	
Consensus	(211)		
		241	270
Leu122-Ser199-Tryp427-Gly431	(241)	GAGATCGTGCTGGAG	MALE ALCOHOLOGICAL COMMENTS CONTROL .
Val127-Asn195-Arg426-Gly431	(241)	GAGATCGTGCTGGAG	Ent. Care Control Control Control of the Care Care Care Care Care Care Care Car
Val120-Thr202-Ile424-Ala433	(241)	GAGATCGTGCTGGAG	
Leu122-Ser199-Arg426-Lys432	(241)	GAGATCGTGCTGGAG	Carrier of the service of the control of the service of the servic
Leu122-Ser199-Arg426-Gly431	(241)	GAGATCSTGCTGGAG	aacgtgaccgagaac
Lys121-Val200-Asn425-Lys432	(241)	GAGATCGTGCTGGAG	unispopulos i limita i versión medicio por unidos esta esta esta esta entre en entre en
Val120-Ile201-Ile424-Ala433	(241)	GAGATCGTGCTGGAG	AACGTGACCGAGAAC
Val120-Ile201B-Ile424-Ala433	(241)	GAGATCGTGCTGGAG	AACGTGACCGAGAAC
Consensus	(241)	GAGATCGŤGCTGGAGA	AACGTGACCGAGAAC
		271	300
Leu122-Ser199-Tryp427-Gly431	(271)	TTCAACATGTGGAAG	aacaacatggt gg ag
Val127-Asn195-Arg426-Gly431	(271)	TTCAACATGTGGAAG	
Val120-Thr202-Ile424-Ala433	(271)	TTCAACATGTGGAAG!	
Leu122-Ser199-Arg426-Lys432	(271)	TTCAACATGTGGAAG	aacaacat ggtgg ag
Leu122-Ser199-Arg426-Gly431	(271)	TTCAACATGTGGAAG	aacaacatggtggag
Lys121-Val200-Asn425-Lys432	(271)	TTCAACATGTGGAAG	AACAACATGGTGGAG
Val120-Ile201-Ile424-Ala433	(271)	TTCAACATGTGGAAGA	AACAACATGGTGGAG
Val120-Ile201B-Ile424-Ala433	(271)	TTCAACATGTGGAAGA	AACAACATGGTGGAG
Consensus	(271)	TTCAACATGTGGAAGA	AACAACATGGTGGAG
		301	330
Leu122-Ser199-Tryp427-Gly431	(301)	CAGATGCACGAGGACA	ATCATCAGCCTGTGG
Val127-Asn195-Arg426-Gly431	(301)	CAGATGCACGAGGAC	
Val120-Thr202-Ile424-Ala433	(301)	CAGATGCACGAGGAC	ATCATCAGCCTGTGG
Leu122-Ser199-Arg426-Lys432	(301)	CAGATGCACGAGGAC	ATCATCAGCCTGTGG
Leu122-Ser199-Arg426-Gly431	(301)	CAGATGCACGAGGAC	ATCATCAGCCTGTGG
Lys121-Val200-Asn425-Lys432	(301)	CAGATGCACGAGGACA	ATCATCAGCCTGTGG
Val120-Ile201-Ile424-Ala433	(301)	CAGATGCACGAGGACA	CONTRACTOR
Val120-Ile201B-Ile424-Ala433	(301)	CAGATGCACGAGGACA	CONTROL TO A DE TRACTO SERVICIONE LA SELECTION
Consensus	(301)	CAGATGCACGAGGACA	AND A THEORY AND A THEORY OF A THEORY OF A STORY OF A S
	/	331	360
Leu122-Ser199-Tryp427-Gly431	(331)	GACCAGAGCCTGAAG	
Val127-Asn195-Arg426-Gly431	(331)	GACCAGAGCCTGAAG	The state was a second
Val120-Thr202-Ile424-Ala433	(331)	GACCAGAGCCTGAAGC	Service Community of the Community of th
	(331)		TOTAL TOTAL TOTAL TOTAL

WO 00/39303	20	/ 65 PCT/US9
	30	7 03
Leu122-Ser199-Arg426-Lys432	(331)	GACCAGAGCCTGAAGCCCTGCGTGAAGCTG
Leu122-Ser199-Arg426-Gly431	(331)	GACCAGAGCCTGAAGCCCTGCGTGAAGCTG
Lys121-Val200-Asn425-Lys432	(331)	GACCAGAGCCTGAAGCCCTGCGTGAA
Val120-Ile201-Ile424-Ala433	(331)	GACCAGAGCCTGAAGCCCTGCGTG
Val120-Ile201B-Ile424-Ala433	(331)	GACCAGAGCCTGAAGCCCTGCGTG
Consensus	(331)	GACCAGAGCCTGAAGCCCTGCGTGAAGCTG
		361 390
Leu122-Ser199-Tryp427-Gly431	(361)	GG
Val127-Asn195-Arg426-Gly431	(361)	ACCCCCTGTGCG TGG GGGCAGGGAACTGC
Val120-Thr202-Ile424-Ala433	(355)	GG
Leu122-Ser199-Arg426-Lys432	(361)	GG
Leu122-Ser199-Arg426-Gly431	(361)	GG
Lys121-Val200-Asn425-Lys432	(357)	
Val120-Ile201-Ile424-Ala433	(355)	
Val120-11e201-11e424-Ala433	(355)	
Consensus	(361)	GG 420
. 100 0 100 T (00 0) (01	1000	391 420
Leul22-Ser199-Tryp427-Gly431	(363)	CAACAGCGTGATCACCCAGGCCTGCCCC
Val127-Asn195-Arg426-Gly431	(391)	AACACCAGCGTGATCACCCAGGCCTGCCCC
Val120-Thr202-Ile424-Ala433	(357)	CGGCGCCACCCAGGCCTGCCCC
Leu122-Ser199-Arg426-Lys432	(363)	CAACAGCGTGATCACCCAGGCCTGCCCC
Leu122-Ser199-Arg426-Gly431	(363)	CAACAGCGTGATCACCCAGGCCTGCCCC
Lys121-Val200-Asn425-Lys432	(359)	CCCCCGTGATCACCCAGGCCTGCCCC
Val120-Ile201-Ile424-Ala433	(355)	GEGGGCATCACCCAGGCCTGCCCC
Val120-Ile201B-Ile424-Ala433	(355)	CCCGGCATCACCCAGGCCTGCCCC
Consensus	(391)	CA CAGCGTGATCACCCAGGCCTGCCCC
	(/	421 450
Leul22-Ser199-Tryp427-Gly431	(391)	AAGGTGAGCTTCGAGCCCATCCCATCCAC
Val127-Asn195-Arg426-Gly431	(421)	AAGGTGAGCTTCGAGCCCATCCCCATCCAC
Val120-Thr202-Ile424-Ala433	(379)	AAGGTGAGCTTCGAGCCCATCCCCATCCAC
Leu122-Ser199-Arg426-Lys432	(391)	AAGGTGAGCTTCGAGCCCATCCAC
Leu122-Ser199-Arg426-Gly431	(391)	AAGGTGAGCTTCGAGCCCATCCCCATCCAC
Lys121-Val200-Asn425-Lys432	(385)	AAGGTGAGCTTCGAGCCCATCCCCATCCAC
Val120-Ile201-Ile424-Ala433	(379)	AAGGTGAGCTTCGAGCCCATCCCCATCCAC
/al120-Ile201B-Ile424-Ala433	(379)	AAGGTGAGCTTCGAGCCCATCCCCCTCCAC
Consensus	(421)	AAGGTGAGCTTCGAGCCCATCCCCATCCAC
		451 480
Leu122-Ser199-Tryp427-Gly431	(421)	TACTGCGCCCCCCCCGCTTCGCCATCCTG
Val127-Asn195-Arg426-Gly431	(451)	TACTGCGCCCCCGCCGGCTTCGCCATCCTG
Val120-Thr202-Ile424-Ala433	(409)	TACTGUECCUECECCGGCTTCGCCATCCTG
Leu122-Ser199-Arg426-Lys432	(421)	TAC : GOCGO COCGGGGTTCGGCCATCCTG
Leu122-Ser199-Arq426-Gly431	(421)	TACTGCGGC CCCCCGGCTTCGCCATCCTG
Lys121-Val200-Asn425-Lys432	(415)	
Val120-Ile201-Ile424-Ala433	(409)	and the second s
Val120-Ile201B-Ile424-Ala433	(409)	
Consensus	(451)	
		481 510
eu122-Ser199-Tryp427-Gly431	(451)	AAGTGCAACGACAAGAAGTTCAACGGCAGC
Val127-Asn195-Arg426-Gly431	(481)	**************************************
Val120-Thr202-Ile424-Ala433	(439)	AAGTGCAACGACAAGAAGTTCAACGGCAGC
Leu122-Ser199-Arg426-Lys432	(451)	AAGTGCAACGACAAGAAGTTCAACGGCAGC
Leu122-Ser199-Arg426-Gly431	(451)	AAGTGCAACGACAAGAAGTTCAACGGCAGC
Lys121-Val200-Asn425-Lys432	(445)	Control of the Contro
Val120-Ile201-Ile424-Ala433	(439)	The state of the s
Val120 11e201 11e424 71e433 Val120 11e201B 11e424 71e433	(439)	the state of the s
Consensus		AAGTGCAACGACAAGAAGTTCAACGGCAGC
Consensus	(40T)	511 540
		540

WO 00/20202	-	/ 65 PCT/US99/31272
WO 00/39303	.31	7 03
Leu122-Ser199-Tryp427-Gly431	(481)	GGCCCTGCACCAACGTGAGCACCGTGCAG
Vall27-Asn195-Arg426-Gly431 Vall20-Thr202-Ile424-Ala433	(511) (469)	GCCCCTGCACCAACGTGAGCACCGTGCAG GGCCCCTGCACCAACGTGAGCACCGTGCAG
		GGCCCTGCACCAACGTGAGCACCGTGCAG GGCCCCTGCACCAACGTGAGCACCGTGCAG
Leu122-Ser199-Arg426-Lys432	(481) (481)	GGCCCCTGCACCAACGTGAGCACCGTGCAG
Leu122-Ser199-Arg426-Gly431	(475)	GGCCCTGCACCAACGTGAGCACCGTGCAG
Lys121-Val200-Asn425-Lys432 Val120-Ile201-Ile424-Ala433	(473)	GGCCCCTGCACCAACGTGAGCACCGTGCAG
Val120-11e201-11e424-Ala433 Val120-11e201B-11e424-Ala433	(469)	GGCCCTGCACCAACGTGAGCACCGTGCAG
Consensus	(511)	GGCCCTGCACCAACGTGAGCACCGTGCAG
Consensus	(311)	541 570
Leu122-Ser199-Tryp427-Gly431	(511)	TGCACCCACGGCATCCGCCCCGTGGTGAGC
Val127-Asn195-Arg426-Gly431	(541)	TGCACCCACGGCATCCGCCCCGTGGTGAGC
Val120-Thr202-Ile424-Ala433	(499)	TGCACCCACGGCATCCGCCCCGTGGTGAGC
Leu122-Ser199-Arg426-Lys432	(511)	TGCACCCACGGCATCCGCCCCGTGGTGAGC
Leu122-Ser199-Arg426-Eys432 Leu122-Ser199-Arg426-Gly431	(511)	TGCACCCACGGCATCCGCCCCGTGGTGAGC
Lys121-Val200-Asn425-Lys432	(505)	TGCACCCACGGCATCCGCCCCGTGGTGAGC
Val120-Ile201-Ile424-Ala433		1.000 000 000 000 000 000 000 000 000 00
Val120-11e201-11e424-Ala433 Val120-11e201B-11e424-Ala433	(499) (499)	TGCACCCACGGCATCCGCCCCGTGGTGAGC TGCACCCACGGCATCCGCCCCGTGGTGAGC
		The Self-Arma Proceeding against State Control of the Control of t
Consensus	(541)	TGCACCCACGGCATCCGCCCCGTGGTGAGC
Inv. 122 Com 100 Emm 427 Cl. 421	(541)	571 600
Leu122-Ser199-Tryp427-Gly431		ACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Val127-Asn195-Arg426-Gly431 Val120-Thr202-Ile424-Ala433	(571) (529)	ACCCAGCTGCTGCTGAACGGCAGCCTGGCC
	, ,	ACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Leu122-Ser199-Arg426-Lys432	(541)	ACCCAGCTGCTGCTGAACGCCAGCCTGGCC
Leu122-Ser199-Arg426-Gly431	(541)	ACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Lys121-Val200-Asn425-Lys432	(535)	ACCCAGCTGCTGCAACGGCAGCCTGGCC
Val120-Ile201-Ile424-Ala433	(529)	ACCCAGCTGCTGAACGCCAGCCTGGCC
Vall20-Ile201B-Ile424-Ala433	(529)	ACCCAGCTGCTGCTGAACGCCAGCCTGGCC
Consensus	(571)	ACCCAGCTGCTGCAACGGCAGCCTGGCC
Leu122-Ser199-Tryp427-Gly431	(571)	630 GAGGAGGCGTGGTGATCCGCAGCGAGAAC
Val127-Asn195-Arg426-Gly431	(601)	GAGGAGGGCGTGGTGATCCGCAGCGAGAAC
Val120-Thr202-Ile424-Ala433	(559)	GAGGAGGCGTGGTGATCCGCAGCGAGAAC
	(571)	GAGGAGGCCTGCTGATCCGCAGCGAGAC
Leu122-Ser199-Arg426-Lys432	(571)	GAGGAGGGCGTGGTGATCCGCAGCGAGAAC
Leu122-Ser199-Arg426-Gly431 Lys121-Val200-Asn425-Lys432	(565)	GAGGAGGCCTGGTGATCCGCAGCGAGAAC
Val120-Ile201-Ile424-Ala433	(559)	GAGGAGGCCTGGTGATCCGCAGCGAGAAC
	(559)	
Val120-Ile201B-Ile424-Ala433		GAGGAGGGCGTGGTGATCCGCAGCGAGAAC
Consensus	(601)	GAGGAGGCGTGGTGATCCGCAGCGAGAAC 631 660
Inv122-Com100 From 427 Clos421	(601)	TTCACCGACAACGCCAAGACCATCATCGTG
Leu122-Ser199-Tryp427-Gly431 Val127-Asn195-Arg426-Gly431		TTCACCGACACGCCAAGACCATCATCGTG
	(631)	そんがちちかめたささと ないのものがからあるとなった。これを大きなないできます。これには、これには、これには、これには、これには、これには、これには、これには、
Val120-Thr202-Ile424-Ala433	(589)	TTCACCGACAACGCCAAGACCATCGTG
Leu122-Ser199-Arg426-Lys432	(601)	TTCACCGACAACGCCAAGACCATCATCGTG
Leu122-Ser199-Arg426-Gly431	(601)	TCACCGACAACGCCAAGACCATCATGGTG
Lys121-Val200-Asn425-Lys432	(595)	TTCACCGACAACGCCAAGACCATCATEGTG
Val120-Ile201-Ile424-Ala433	(589)	TEACCGACAACGCCAAGACCATCATCGTG
Val120-Ile201B-Ile424-Ala433	(589)	TTCRCCGACAACGCCAAGACCATCATGGTG
Consensus	(631)	TTCACCGACAACGCCAAGACCATCATCGTG
Toul 22 - Cov. 1 00 m 427 421	16211	661 690
Leu122-Ser199-Tryp427-Gly431	(631)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC
Val127-Asn195-Arg426-Gly431	(661)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC
Val120-Thr202-Ile424-Ala433	(619)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC
Leu122-Ser199-Arg426-Lys432	(631)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC
Leu122-Ser199-Arg426-Gly431	(631)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC
Lys121-Val200-Asn425-Lys432	(625)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC
Val120-Ile201-Ile424-Ala433	(619)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC

WO 00/39303	32	/ 65	PCT/US99/31272
Val120-11e201B-11e424-Ala433	(619)		TCAACTGC
Consensus	(661)	CAGCTGAAGGAGAGCGTGGAGA	and the stay of th
		691	720
Leu122-Ser199-Tryp427-Gly431	(661)	ACCCCCCCAACAACAACACCC	GCAAGAGC
Val127-Asn195-Arg426-Gly431	(691)	ACCCGCCCCAACAACAACACCC	GC AAGAG C
Val120-Thr202-Ile424-Ala433	(649)		17.3 F Mark State 2002/2004 - 17.
Leu122-Ser199-Arg426-Lys432	(661)	ACCCCCCCAACAACAACACCC	10 m 2 m 4 m 5 m 5 m 5 m 5 m 5 m 5 m 5 m 5 m 5
Leu122-Ser199-Arg426-Gly431	(661)	ACCCCCCCAACAACAACACCC	Later Andrew College College College
Lys121-Val200-Asn425-Lys432	(655)	ACCCGCCCCAACAACACACCC	The Copy of Actions, 1
Val120-Ile201-Ile424-Ala433	(649)	To an a Company and a Service of a service property of the service and a service property of the service of the	The first of the second of the first of the second of the
Vall20-Ile201B-Ile424-Ala433	(649)	ACCOCCCAACAACAACACCC	1. BOLD 2.115 4
Consensus	(691)	ACCCGCCCCAACAACAACACCCC	750
Leu122-Ser199-Tryp427-Gly431	(691)	ATCACCATCGGCCCCGGCCGCG	
Val127-Asn195-Arg426-Gly431	(721)	ATCACCATCGGCCCGGCCGCG	-11 Johnson 1000 600 600 11 F
Val120-Thr202-Ile424-Ala433	(679)	ATCACCATCGGCCCGGCCGCG	5 - actor definition describe
Leu122-Ser199-Arg426-Lys432	(691)	ATCACCATCGGCCCGGCCGCG	AL A CATALON & A ACRESCA A. A. A.
Leu122-Ser199-Arg426-Gly431	(691)	ATCACCATCGGCCCCGGCCGCG	7. P. Branch Co. (1975)
Lys121-Val200-Asn425-Lys432	(685)	The state of the s	** ** *** *** *** *** *** *** *** ***
Val120-Ile201-Ile424-Ala433	(679)	FEX.03040400 \$5000 \$5000 \$11 APPENDING \$12 APPENDING \$12 APPENDING \$12 APPENDING \$13 A	Province and an artifacture of the control of the c
Vall20-Tle201B-Ile424-Ala433	(679)	ATCACCATCGGCCCCGGCCGCG	CCTTCTAC
Consensus	(721)	ATCACCATCGGCCCGGCCGCG	CCTTCTAC
		751	780
Leu122-Ser199-Tryp427-Gly431	(721)	GCCACCGGCGACATCATCGGCG.	4 5 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4
Val127-Asn195-Arg426-Gly431	(751)	GCCACCGGCGACATCATCGGCG	ACATCCGC
Val120-Thr202-Ile424-Ala433	(709)	GCCACCGGCGACATCATCGGCG	AND THE SECOND PROPERTY OF THE SECOND PROPERT
Leul22-Ser199-Arg426-Lys432	(721)	GCCACCGGCGACATCATCGGCG	A Paragraphy of the Control of the C
Leul22-Ser199-Arg426-Gly431	(721)	GCCACCGGCGACATCATCGGCG	200 - 200 -
Lys121-Val200-Asn425-Lys432	(715)	GCCACCGGCGACATCATCGGCG	e siene redicididenmen.
Val120-Ile201-Ile424-Ala433	(709)	GCCACCGGCGACATCATCGGCG	A DE MARKET PER PARAMETER
Vall20-Ile201B-Ile424-Ala433	(709)	GCCACGGCGACATCATCGGCG	THE RESIDENCE OF THE PROPERTY OF THE PERSON
Consensus	(751)	GCCACCGGCGACATCATCGGCGA	RCATCCGC 810
Leu122-Ser199-Tryp427-Gly431	(751)	CAGGCCCACTGCAACATCAGCG	
Val127-Asn195-Arg426-Gly431	(781)	CAGGCCCACTGCAACATCAGCG	nadd common programma and common or
Val120-Thr202-Ile424-Ala433	(739)	CAGGCCCACTGCAACATCAGCG	
Leu122-Ser199-Arg426-Lys432	(751)	CAGGCCCACTGCAACATCAGCG	Commence of the Commence of th
Leu122-Ser199-Arg426-Gly431	(751)	CAGGCCCACTGCAACATCAGCG	
Lys121-Val200-Asn425-Lys432	(745)	CAGGCCCACTGCAACATCAGCG	nergy American Control of the American Control of the Control of t
Val120-Ile201-Ile424-Ala433	(739)	CAGGCCCACTGCAACATCAGCG	Value - 1500 (1940), C. L. Marie (1940), Species Communication (1940)
Val120-Ile201B-Ile424-Ala433	(739)	CAGGCCCACTGCAACATCAGCG	
Consensus	(781)	CAGGCCCACTGCAACATCAGCG	transmitted the debag and a second
		811	840
Leu122-Ser199-Tryp427-Gly431	(781)	TGGAACAACACCCTGAAGCAGA!	male and market market and the second control of the control of th
Val127-Asn195-Arg426-Gly431	(811)	TGGAACAACACCCTGAAGCAGA	
Val120-Thr202-Ile424-Ala433	(769)	TGGAACAACACCCTGAAGCAGA'	CONTRACTOR
Leu122-Ser199-Arg426-Lys432	(781)	TGGAACAACACCCTGAAGCAGA	Contract Contract of the Contr
Leu122-Ser199-Arg426-Gly431	(781)	TGGAACAACACCCTGAAGCAGA	
Lys121-Val200-Asn425-Lys432	(775)	TGGAACAACACCCTGAAGCAGA	
Val120-Ile201-Ile424-Ala433	(769)	TGGAACAACACCCTGAAGCAGA	17 C 17 17 17 17 17 17 17 17 17 17 17 17 17
Val120-Ile201B-Ile424-Ala433	(769)	TGGAACAACACCCTGAAGCAGA	CP AND COLOR OF CONTRACTORS
Consensus	(811)	TGGAACAACACCCTGAAGCAGA	
Leu122-Ser199-Trum427-Clu421	/ Q 1 1 v	NACCTCC SCCCC ACTTCCCCA	870 27767CC
Leu122-Ser199-Tryp427-Gly431 Val127-Asn195-Arg426-Gly431	(841)	AAGCTGCAGGCCCAGTTCGGCAAAGCTGCAGGCCCAGTTCGGCAA	The state of the s
Val120-Thr202-Ile424-Ala433	(799)	AAGCTGCAGGCCCAGTTCGGCA	WASHINGTON TO SEE AND
Leu122-Ser199-Arg426-Lys432	(811)	AAGCTGCAGGCCCAGTTCGGCAZ	
Jorijo mysasz	(011)	Land to the of the of the other than	· The Marine State Co

WO 00/39303	33	/ 65 PCT/US99/31272
Leu122-Ser199-Arg426-Gly431	(811)	AAGCTGCAGGCCCAGTTCGGCAACAAGACC
Lys121-Va1200-Asn425-Lys432	(805)	AAGCTGCAGGCCCAGTTCGGCAACAAGACC
Val120-Ile201-Ile424-Ala433	(799)	AAGCTGCAGGCCCAGTTCGGCAACAAGACC
Val120-Ile201B-Ile424-Ala433	(799)	AAGCTGCAGGCCCAGTTCGGCAACAAGACC
Consensus	(841)	
		871 900
Leu122-Ser199-Tryp427-Gly431	(841)	ATCGTGTTCAAGCAGAGCAGCGGCGGCGAC
Val127-Asn195-Arg426-Gly431	(871)	
Val120-Thr202-Ile424-Ala433	(829)	ATCGTGTTCAAGCAGAGCAGCGGCGGCGAC
Leu122-Ser199-Arg426-Lys432	(841)	
Leu122-Ser199-Arg426-Gly431	(841)	
Lys121-Val200-Asn425-Lys432	(835)	ATCGTGTTCAAGCAGAGCAGCGGCGGCGAC
Val120-Ile201-Ile424-Ala433	(829)	ATCGTGTTCAAGCAGAGCAGCGGCGCGAC
Val120-Ile201B-Ile424-Ala433	(829)	
Consensus	(871)	
		901 930
Leu122-Ser199-Tryp427-Gly431	(871)	CCCGAGATCGTGATGCACAGCTTCAACTGC
Val127-Asn195-Arg426-Gly431	(901)	CCCGAGATCGTGATGCACAGCTTCAACTGC
Val120-Thr202-Ile424-Ala433	(859)	CCCGAGATCGTGATGCACAGCTTCAACTGC
Leu122-Ser199-Arg426-Lys432	(871)	CCCGAGATCGTGATGCACAGCTTCAACTGC
Leu122-Ser199-Arg426-Gly431	(871)	CCCGAGATCGTGATGCACAGCTTCAACTGC
Lys121-Val200-Asn425-Lys432	(865)	CCCGAGATCGTGATGCACAGCTTCAACTGC
Val120-Ile201-Ile424-Ala433	(859)	CCCGAGATCGTGATGCACAGCTTCAACTGC
Vall20-Ile201B-Ile424-Ala433	(859)	CCCGAGATCGTGATGCACAGCTTCAACTGC
Consensus	(901)	
Combabas	(501)	931 960
Leu122-Ser199-Tryp427-Gly431	(901)	GGCGGCGAGTTCTTCTACTGCAACAGCACC
Val127-Asn195-Arg426-Gly431	(931)	GGCGGCGAGTTCTTCTACTGCAACAGCACC
Val120-Thr202-Ile424-Ala433	(889)	GGCGGCGAGTICTTCTACTGCAACACCACC
Leu122-Ser199-Arg426-Lys432	(901)	GGCGGCGAGTTCTTCTACTGCAACAGCACC
Leu122-Ser199-Arg426-Gly431	(901)	GGCGGCGAGTTCTTCTACTGCAACAGCACC
Lys121-Val200-Asn425-Lys432	(895)	GGCGGCGAGTTCTTCTACTGCAACAGCACC
Val120-Ile201-Ile424-Ala433	(889)	GGCGGCGAGTTCTTCTACTGCAACAGCACC
Val120-Ile201B-Ile424-Ala433	(889)	GGCGCCAGTICTTCTACTGCAACAGCACC
Consensus	(931)	GGCGGCGAGTTCTTCTACTGCAACAGCACC
oonsensus	()))	961 990
Leu122-Ser199-Tryp427-Gly431	(931)	CAGCTGTTCAACAGCACCTGGAACAACACC
Val127-Asn195-Arg426-Gly431	(961)	CAGCTGTTCAACAGCACCTGGAACAACACC
Val120-Thr202-Ile424-Ala433	(919)	CAGCTGTTCAACAGCACCTGGAACAACACC
Leu122-Ser199-Arg426-Lys432	(931)	CAGCTGTTCAACAGCACCTGGAACACACC
Leu122-Ser199-Arg426-Gly431	(931)	CAGCIGITCAACAGCACCTGGAACAACACC
Lys121-Val200-Asn425-Lys432	(925)	CAGCTGTTCAACAGCACCTGGAACACACC
Val120-Ile201-Ile424-Ala433		CAGCTGTTCAACAGCACCTGGAACAACACC
Vall20-Ile201B-Ile424-Ala433		CAGCTGTTCAACAGCACCTGGAACAACACC
Consensus		CAGCTGTTCAACAGCACCTGGAACAACACC
3030348	(301)	991 1020
Leu122-Ser199-Tryp427-Gly431	(961)	ATCGGCCCCAACAACACCAACGGCACCATC
Val127-Asn195-Arg426-Gly431	(991)	ATCGGCCCCAACAACACCAACGGCACCATC
Val120-Thr202-Ile424-Ala433	(949)	ATCGGCCCCAACACACCCAACGGCACGATC
Leu122-Ser199-Arg426-Lys432	(961)	The state of the s
Leu122-Ser199-Arg426-Gly431	(961)	ATCGGCCCAACACCCCAACGGCACCATC
Lys121-Val200-Asn425-Lys432	(955)	ATCGGCCCAACACACCAACGGCACCATC
Val120-Ile201-Ile424-Ala433		ATCGGCCCCAACACACCGCACCATC
Val120 I1c201 I1c424 Ala433	(949)	ATCGGCCCAACAACACCAACGGCACCATC
Consensus		ATCGGCCCAACACACCCAACGGCACCATC
Consensus	(221)	1021 1050
Leu122-Ser199-Tryp427-Gly431	(991)	ACCCTGCCCTGCCGCATCAAGCAGATCATC
corray irypazi ciyadi	(221)	23COOTSCOOTSCOOTSCOOTSCOOTSCOOTSCOOTSCOOT

```
Val127-Asn195-Arg426-Gly431
                              (1021) ACCCTGCCCTGCCGCATCAAGCAGATCATC
 Val120-Thr202-Ile424-Ala433
                               (979) ACCCTGCCTGCCGCATCAAGCAGATCATC
 Leu122-Ser199-Arg426-Lys432
                               (991) ACCCTGCCCTGCCGCATCAGCAGCATCATC
 Leu122-Ser199-Arg426-Gly431
                               (991) ACCCTGCCTGCCGCATCAAGCAGATCATC
 Lys121-Val200-Asn425-Lys432
                               (985) ACCCTGCCCTGCCGCATCAAGCAGATCATC
 Val120-Ile201-Ile424-Ala433
                               (979) ACCCTGCCCTGCCGCATCAAGCAGATCATC
Val120-Ile201B-Ile424-Ala433
                               (979) ACCCTGCCCTGCCGCATCAAGCAGATCATC
                   Consensus
                              (1021) ACCCTGCCCTGCCGCATCAAGCAGATCATC
                              (1021) AACCGCTGGGGCGCAAGGCCATGTACGCC
Leu122-Ser199 Tryp427-Gly431
Val127-Asn195-Arg426-Gly431
                              (1051) AACCGCGGCGGCGCAAGGCCATGTACGCC
Val120-Thr202-Ile424-Ala433
                              (1009) -----GGCGGC---GCCATGTACGCC
                              (1021) AACCGCGGCGGCAACAAGGCCATGTACGCC
Leu122-Ser199-Arg426-Lys432
Leu122-Ser199-Arg426-Glv431
                              (1021) AACCGCGGCAGCGCAAGGCCATGTACGCC
                              (1015) AAC-----GCCCCCAAGGCCATGTACGCC
Lys121-Val200-Asn425-Lys432
 Val120-Ile201-Ile424-Ala433
                              (1009) -----GGCGGC---GCCATGTACGCC
                              (1009) -----GGCGGC---GCCATGTACGCC
Val120-Ile201B-Ile424-Ala433
                   Consensus
                              (1051) AACCGC G GGCGGCAAGGCCATGTACGCC
                                     1081
                              (1051) CCCCCCATCCGCGGCCAGATCCGCTGCAGC
Leu122-Ser199 Tryp427-Gly431
 Val127-Asn195-Arg426-Gly431
                              (1081) CCCCCATCCGCGGCCAGATCCGCTGCAGC
 Val120-Thr202-Ile424-Ala433
                              (1027) CCCCCCATCCGCGGCCAGATCCGCTGCAGC
                              (1051) CCCCCATCCGCGCCAGATCCGCTGCAGC
Leu122-Ser199-Arg426-Lys432
Leu122-Ser199-Arg426-Gly431
                              (1051) CCCCCCATCCGCGCCAGATCCGCTGCAGC
Lys121-Val200-Asn425-Lys432
                              (1039) CCCCCCATCCGCGGCCAGATCCGGTGCAGC
 Val120-Ile201-Ile424-Ala433
                              (1027) CCCCCCATCCGCGCCAGATCCGCTGCAGC
Val120-Ile201B-Ile424-Ala433
                              (1027) CCCCCCATCCGCGCCAGATCCGCTGCAGC
                              (1081) CCCCCCATCCGCGGCCAGATCCGCTGCAGC
                   Consensus
                                     1111
                              (1081) AGCAACATCACCGGCCTGCTGCTGACCCGC
Leu122-Ser199 Tryp427-Gly431
Val127-Asn195-Arg426-Gly431
                              (1111) AGCAACATCACCGGCCTGCTGCTGACCCGC
Val120-Thr202-Ile424-Ala433
                              (1057) AGCARCATCACCGGCCTGCTGACCCGC
Leu122-Ser199-Arg426-Lys432
                              (1081) AGCAACATCACCGGCCTCCTGCTGACCCGC
Leu122-Ser199-Arg426-Gly431
                              (1081) AGCARCATCACCGGCCTGCTGACCCGC
                              (1069) AGCAACATCACCGGCCTGCTGCTGACCCGC
Lys121-Val200-Asn425-Lys432
                              (1057) AGCRACATCAGGGGCCTGGTGGTGGTGAGGGGC
Val120-Ile201-Ile424-Ala433
                              (1057) AGGAACATCACCGGCCTGCTGCTGACCCGC
Val120-Ile201B-Ile424-Ala433
                   Consensus
                              (1111) AGCAACATCACCGGCCTGCTGACCCGC
                              (1111) GACGOCGGCAAGGAGATCAGCAACACCACC
Leu122-Ser199 Tryp427-Gly431
Val127-Asn195-Arg426-Gly431
                              (1141) GACGGCGGCAAGGAGATCAGCAACCACC
Val120-Thr202-Ile424-Ala433
                              (1087) GACGGCGCAAGGAGATCAGCAACACCACC
Leu122-Ser199-Arg426-Lys432
                              (1111) GACGGCGGCAAGGAGATCAGCAACACCACC
Leu122-Ser199-Arg426-Gly431
                              (1111) GACGCCGCCRAGGAGATCAGCAACACCACC
Lys121-Val200-Asn425-Lys432
                              (1099) GACGGCGGCAAGGAGATCAGCAACACCACC
Val120-Ile201-Ile424-Ala433
                              (1087) GACGGCGCAAGGAGATCAGCAACACCACC
Val120-Ile201B-Ile424-Ala433
                              (1087) GACGGCGGCAAGGAGATCAGCAACACCACC
                   Consensus
                              (1141) GACGGCGGCAAGGAGATCAGCAACACCACC
                                     1171
                                                               1200
Leu122-Ser199 Tryp427-Gly431
                              (1141) GAGATCTTCCGCCCCGGCGGCGCGACATG
Val127-Asn195-Arg426-Gly431
                              (1171) GAGATCTTCCGCCCCGGCGCGCGACATG
Val120-Thr202-Ile424-Ala433
                              (1117) GAGATCTTCCGCCCGGCGGCGCGACATG
Leu122-Ser199-Arg426-Lys432
                              (1141) GAGATCTTCCGCCCCGGCGCGCGCGACATG
                              (1141) GAGATCTTCCGCCCGGCGGCGGCGACATG
Leu122-Ser199-Arg426-Gly431
Lys121-Val200-Asn425-Lys432
                              (1129) GAGATCTTCCGCCCCGGCGGCGCGACATG
Val120-Ile201-Ile424-Ala433
                              (1117) GAGATCTTCCGCCCCGGCGCGCGCGACATG
Val120-Ile201B-Ile424-Ala433
                              (1117) GAGATCTTCCGCCCCGGCGGCGGCGACATG
```

Consensus	(1171)	GAGATCTTCCGCCCCGGCGGCGGCGACATG
Consensus	(11/1/	
Tou 122 Car 100 Trum 427 Clar421	(1171)	1200
Leu122-Ser199 Tryp427-Gly431	(1171)	CGCGACAACTIGCSCAGCIACCTGTACAAG
Val127-Asn195-Arg426-Gly431	(1201)	CGCGECAACTSSC CAGCGAGCTGTACAAG
Val120-Thr202-Ile424-Ala433	(1147)	COCGACARET ESCACAGO GAGOTTA A AAG COCGACAROS GAGOTTA AAG
Leu122-Ser199-Arg426-Lys432	(1171)	
Leu122-Ser199-Arg426-Gly431	(1171)	COCCACA ACTION CONTRACTOR ACAM
Lys121-Val200-Asn425-Lys432	(1159)	CCCCACAACTCCCCCCCCCCCCCCCCTCTACAAC
Val120-Ile201-Ile424-Ala433	(1147)	COCGACAACTOCCGCCACCTCTACAAG
Val120-Ile201B-Ile424-Ala433	(1147)	GACAACIGG SCACCIACITATA AG
Consensus	(1201)	CGCGACAACTGGCGCGAGCTGTACAAG
	, ,	1231 1260
Leu122-Ser199 Tryp427-Gly431	(1201)	TACAAGGIGGIGAAGATCGAGCCCCTGGGC
Val127-Asn195-Arg426-Gly431	(1231)	TACA GGT GG GAACATCCACCC TGGGC
Val120-Thr202-Ile424-Ala433	(1231)	9.5
Leu122-Ser199-Arg426-Lys432		TACAAGGTEE GAAGAT GAGCC TGGGC
	(1201)	TACAAGGIGGIGAAGATCGAGCCCCTGGGC
Leu122-Ser199-Arg426-Gly431	(1201)	TACAAGGTGGTGAAGATCGAGCCCCTGGGC
Lys121-Val200-Asn425-Lys432	(1189)	TACAAGGT <u>G</u> GGAAGATC <u>GA</u> GTCCCTGGG
Val120-Ile201-Ile424-Ala433	(1177)	TACAAGGTGGTGAAGATCGAGCCCCTGGGC
Val120-Ile201B-Ile424-Ala433	(1177)	TACAAGGTGGTGAAGATCGAGCCCCTGGGC
Consensus	(1231)	TACAAGGTGGTGAAGATCGAGCCCCTGGGC
		1261 1290
Leu122-Ser199 Tryp427-Gly431	(1231)	GTGGCCCCACCAAGGCCAAGGGCGCGTG
Val127-Asn195-Arg426-Gly431	(1261)	CTGGCCCCACCAAGGCCAAGGGCCGCGTG
Val120-Thr202-Ile424-Ala433	(1207)	CTGGGCCCGACCAAGGCCCAAGGCCCGGGGTG
Leu122-Ser199-Arg426-Lys432	(1231)	GTGGCCCCACCAAGGCCAAGGCGCGCGTG
Leu122-Ser199-Arg426-Gly431	(1231)	
Lys121-Val200-Asn425-Lys432	(1231) (1219)	GTGGCCCCACCAAGCCCAAGCGCCCCTG
Val120-Ile201-Ile424-Ala433		GTGGCCCCCACCAAGGCCAAGGCCGCCCTG
	(1207)	GTGGCCCCCACCAAGGCCAAGGGCCGCGTG
Val120-Ile201B-Ile424-Ala433	(1207)	GTGGCCCCACCAAGGCCAAGGGCCGCGTG
Consensus	(1261)	GTGGCCCCACCAAGGCCAAGCGCCGCGTG
		1291 1320
Leu122-Ser199 Tryp427-Gly431	(1261)	GTGCAGCGCGAGAAGCGCGCCGTGACCCTG
Val127-Asn195-Arg426-Gly431	(1291)	GTGCAGCGCGAGAAGCGCGCCCTGACCCTG
Val120-Thr202-Ile424-Ala433	(1237)	GTGGAGCGCGAGAAGCGCGCGCTGACCCTG
Leu122-Ser199-Arg426-Lys432	(1261)	G G A/AC G G G G
Leu122-Ser199-Arg426-Gly431	(1261)	GTCCAGCGCCAGAAGCGCGCGGTCAGCGTG
Lys121-Val200-Asn425-Lys432	(1249)	GIGCAGCOGAGNAGCGC COGTGACCCTG
Val120-Ile201-Ile424-Ala433	(1237)	GIGCAGCGGGAGAAGGGCCCCCTGACCCCTG
Val120-Ile201B-Ile424-Ala433	(1237)	GTGCAGCGCGAGAAGCGCGCCCTGACCCTG
Consensus	(1291)	
Consensus	(1291)	GTGCAGCGCGAGAAGCGCGCCGTGACCCTG
Tau100 0 100 m 407 01 401	(1001)	1321 - 1350
Leu122-Ser199 Tryp427-Gly431	(1291)	GGCGC A YOTH OF TGGGC TCCTGGCCC
Val127-Asn195-Arg426-Gly431	(1321)	GGCGCCATGTTCCTGGGCCCC
Val120-Thr202-Ile424-Ala433	(1267)	GCCGCCATICITYCCTGG
Leu122-Ser199-Arg426-Lys432	(1291)	GCCGCCATGTTCCTIGGGCCTTCCTTGCGCGC
Leu122-Ser199-Arg426-Gly431	(1291)	GGCGCCATCTTCCTGGGCTTCCTGGGCGC
Lys121-Val200-Asn425-Lys432	(1279)	GGCGCCATGTTCCTGGGGTTCCTCGGCCCC
Val120-Ile201-Ile424-Ala433	(1267)	GGCCCCATGTTC TGG CTTCGT CSCCC
Val120-Ile201B-Ile424-Ala433	(1267)	GGCGCCATGTTCCTGGGCTTCCTGGGCGCC
Consensus	(1321)	GGCGCCATGTTCCTGGGCTTCCTGGGCGCC
Consciisus	(+021)	1351 1380
Leu122-Ser199 Tryp427-Gly431	(1321)	SCGGCA CA CAIGGGGGGGGGGGGAGGGTG
Val127-Asn195-Arg426-Gly431	(1351)	
Val120-Thr202-Ile424-Ala433		$C_{A} = C_{A} = C_{A} = C_{A}$
	(1297)	GCCGGCAGCACCATGGGCCCCCGCAGCCTG
Leu122-Ser199-Arg426-Lys432	(1321)	GCCGCAGCACCATGGGCCCCCCCAGCCTG
Leu122-Ser199-Arg426-Gly431	(1321)	GCCGGCAGCACCATGGGCGCCCCGTAGCCTG

WO 00/39303	36 /	65 PCT/US99/31272
Lys121-Val200-Asn425-Lys432	(1309)	GCCGGCAGCACCATGGGCGCCCGCAGCCTG
Val120-Ile201-Ile424-Ala433	(1297)	The state of the s
Val120-Ile201B-Ile424-Ala433	(1297)	GCCGCAGCACCATGGGCGCCCGCAGCCTG
	(1351)	GCCGCAGCACCATGGGCGCCCGCAGCCTG
Consensus	(1221)	1381 1410
1 1 2 2 C 1 0 0 M 4 2 7 C 1 4 2 1	/12E1V	——————————————————————————————————————
Leu122-Ser199 Tryp427-Gly431	(1351)	ACCCTGACCGTGCAGGCCGCCAGCTGCTG
Val127-Asn195-Arg426-Gly431	(1381)	ACCCTGACCGTGCAGGCCCGCCAGCTGCTG
Vall20-Thr202-Ile424-Ala433	(1327)	ACCCTGACCGTGCAGGCCCGCCAGCTGCTG
Leu122-Ser199-Arg426-Lys432	(1351)	ACCCTGACCGTGCAGGCCGCCAGCTGCTG
Leu122-Ser199-Arg426-Gly431	(1351)	ACCCTGACCGTGCAGCCCGCCAGCTGCTG
Lys121-Val200-Asn425-Lys432	(1339)	ACCCTGACCGTGCAGCCCGCCAGCTGCTG
Val120-Ile201-Ile424-Ala433	(1327)	ACCCTGACCGTGCAGCCCGCCAGCTGCTG
Val120-Ile201B-Ile424-Ala433	(1327)	ACCCTGACCGTGCAGGCCCGCCAGCTGCTG
Consensus	(1381)	ACCCTGACCGTGCAGGCCCGCCAGCTGCTG
		1411 1440
Leu122-Ser199 Tryp427-Gly431	(1381)	AGCGGCATCGTGCAGCAGCAGCAACCTG
Val127-Asn195-Arg426-Gly431	(1411)	AGCGGCATCGTGCAGCAGCAGAACAACCTG
Val120-Thr202-Ile424-Ala433	(1357)	AGCGGCATCGTGCAGCAGCAGCAACCTG
Leu122-Ser199-Arg426-Lys432	(1381)	AGCGGCATCGTGCAGCAGCAGAACAACCTG
Leu122-Ser199-Arg426-Gly431	(1381)	AGCGGCATCGTGCAGCAGCAGAACAACCTG
Lys121-Val200-Asn425-Lys432	(1369)	AGCGCCATCGTGCAGCAGCAGCAACCTG
Val120-Ile201-Ile424-Ala433	(1357)	AGCGGCATCGTGCAGCAGCAGCAACAACCTG
Val120-Ile201B-Ile424-Ala433	(1357)	AGCGCCATCGTGCAGCAGCAGCAACCAGCTG
Consensus	(1411)	AGCGGCATCGTGCAGCAGCAGAACAACCTG
Consensus	(1111/	1441 1470
Leu122-Ser199 Tryp427-Gly431	(1411)	CTGCGCGCCATCGAGGCCCCAGCACCACCTG
Val127-Asn195-Arg426-Gly431	(1441)	CTGCGCGCCATCGAGGCCCAGCAGCACCTG
Val120-Thr202-Ile424-Ala433	(1387)	CTGCGCGCCATCGAGGCCCCAGCAGCACCTG
Leu122-Ser199-Arg426-Lys432	(1411)	CTGCGCGCCATCGAGGCCCCAGCAGCACCTG
Leu122-Ser199-Arg426-Eys432	(1411)	CTGCGCGCCATCGAGGCCCAGCAGCACCTG
	(1399)	CTGCGCGCCATCGAGGCCCAGCAGCACCTG
Lys121-Val200-Asn425-Lys432		CTGCGCGCCATCGAGGCCCAGCAGCACCTG
Val120-Ile201-Ile424-Ala433	(1387)	 S S. Aggrego C. Tree (C. Cresting Control of Contr
Val120-Ile201B-Ile424-Ala433	(1387)	CTGCGCGCCATCGAGGCCCCAGCAGCACCTG
Consensus	(1441)	CTGCGCGCCATCGAGGCCCAGCACCTG
		1471 1500
Leu122-Ser199 Tryp427-Gly431	(1441)	CTECAGCTGACCETGTGGGGCATCAAGCAG
Val127-Asn195-Arg426-Gly431	(1471)	CTGCAGCTGACCGTGTGGGGCATCAAGCAG
Val120-Thr202-Ile424-Ala433	(1417)	CTGCAGCTGACCGTGTGGGGCATCAAGCAG
Leu122-Ser199-Arg426-Lys432	(1441)	CTGCAGCTGACCGTGTGGGGCATCAAGCAG
Leu122-Ser199-Arg426-Gly431	(1441)	CTGCAGCTGACCGTGTGGGGCATCAAGCAG
Lys121-Va1200-Asn425-Lys 4 32	(1429)	CTGCAGCTGACCGTGTGGGGCATCAAGGAG
Val120-Ile201-Ile424-Ala433	(1417)	CTGCAGCTGACCGTGTGGGGCATCAAGCAG
Vall20-Ile201B-Ile424-Ala433	(1417)	CTGCAGCTGACCGTGTGGGGCATCAAGCAG
Consensus	(1471)	CTGCAGCTGACCGTGTGGGGCATCAAGCAG
		1501 1530
Leu122-Ser199 Tryp427-Gly431	(1471)	GTGCAGGCCGGGGTGCTGGCCGTGGAGCGC
Val127-Asn195-Arg426-Gly431	(1501)	CTGCAGGCCCGCGTGCTGGCCGTGGAGCGC
Val120-Thr202-Ile424-Ala433	(1447)	CTGCAGGCCCGCGTGCTGGCCCTGGAGCGC
Leu122-Ser199-Arg426-Lys432		CHGGAGGGGAGGGTHGGAGGGC
Leu122-Ser199-Arg426-Gly431		CTGCAGGCCGCGTGCTGGCCGTGGAGGGC
Lys121-Val200-Asn425-Lys432		CTGCAGCCCGCGTGCTGGCCGTGGAGCGC
Val120-Ile201-Ile424-Ala433	(1447)	CTGCAGGCCCGCGTGCTGCCGTGGAGCGC
Val120 Ile201 Ile424 Ala433		CTGCAGGCCCGCGTGCTGGCCGTGGAGCGC
Consensus		CTGCAGGCCCGCGTGCTGGCCGTGGAGCGC
Consensus	(+00+)	1531 1560
Leu122-Ser199 Tryp427-Gly431	(1501)	TACCTGAAGGACCAGCAGCTGCTGGGCATC
Vall27-Asn195-Arg426-Gly431	(1531)	A CONTRACTOR OF THE PROPERTY O
va112/-A511133-A19420-G19431	(1777)	110013000000000000000000000000000000000

65

WO 00/39303	38	65	PCT/US99/31272
		1711	1740
Leu122-Ser199 Tryp427-Gly431	(1681)	ATCGAGGAG	AGCCAGAACCAGCAGGAGAAG
Val127-Asn195-Arg426-Gly431	(1711)	ATCGAGGAG	AGCCAGAACCAGCAGGAGAAG
Val120-Thr202-Ile424-Ala433	(1657)		AGCCAGAACCAGCAGGAGAAG
Leu122-Ser199-Arg426-Lys432	(1681)		AGCCAGAACCAGCAGGAGAAG
Leu122-Ser199-Arg426-Gly431	(1681)	Author Continued the suit of the continued to	AGCCAGAACCAGCAGGAGAAG
Lys121-Val200-Asn425-Lys432	(1669)	The state of the s	AGCCAGAACCAGCAGGAGAAG
Val120-Ile201-Ile424-Ala433	(1657)	45-2015. 7-4-2-4-2-2-2-2-2-2-2-2-2-2-2-2-2-2-2-2-	AGCCAGAACCAGCAGAGAAG
Val120-Ile201B-Ile424-Ala433	(1657)	County Comments of the Comment	ĀĢCCAGAACCAGCAGGAGAAG
Consensus		A mount becomes understanding the first control designation	AGCCAGAACCAGCAGGAGAAG
consensus	(1/11/	1741	1770
Leu122-Ser199 Tryp427-Gly431	(1711)		
2.2 2	(1741)		GAGCTGCTGGACTAG
Val127-Asn195-Arg426-Gly431		17 1 1 1 2 2 4 1 1 1 1 1 1 1 1 1 1 1 1 1 1	GAGCTGCTGGAGCAAG
Val120-Thr202-Ile424-Ala433	(1687)		GAGCTGCTGGAGCAAG
Leu122-Ser199-Arg426-Lys432	(1711)	25/5 To Annual Control (1995)	GAGCTGCTGGAGCAAG
Leu122-Ser199-Arg426-Gly431	(1711)	ACTUAL TO CONCURSIONAL CONTRACTOR	GAGCTGCTGGAGCAAG
Lys121-Val200-Asn425-Lys432	(1699)		GAGCTGCTGGAGCAAG
Val120-Ile201-Ile424-Ala433	(1687)		GAGCTGCTGGAGCAAG
Vall20-Ile201B-Ile424-Ala433	(1687)		GAGCTGCTGGAGCAAG
Consensus	(1741)	AACGAGCAG	GAGCTGCTGGAGCAAG
		1771	1800
Leu122-Ser199 Tryp427-Gly431	(1741)	TGGGCCAGC	CTGTGGAACTGGTTCGACATC
Val127-Asn195-Arg426-Gly431	(1771)		CTGTGGAACTGGTTCGACATC
Val120-Thr202-Ile424-Ala433	(1717)		CTGTGGAACTGGTTCGACATC
Leu122-Ser199-Arg426-Lys432	(1741)	A	CTGTGGAACTGGTTCGACATC
Leu122-Ser199-Arg426-Gly431	(1741)		CTGTGGAACTGGTTCGACATC
Lys121-Val200-Asn425-Lys432	(1729)	CONTRACTOR SERVICE CONTRACTOR SERVICES	CTGTGGAACTGGTTCGACATC
Val120-Ile201-Ile424-Ala433	(1717)		CTGTGGAACTGGTTCGACATC
Val120-Ile201B-Ile424-Ala433	(1717)		CTGTGGAACTGGTTCGACATC
Consensus	(1771)	Court was right to a become more convenient to the Tarketon	CTGTGGAACTGGTTCGACATC
consensus	(1,11)	1801	1830
Leu122-Ser199 Tryp427-Gly431	(1771)		CTGTGGTACATCAAGATCTTC
Val127-Asn195-Arg426-Gly431	•		
	(1801)		CTGTGGTACATCAAGATCTTC
Val120-Thr202-Ile424-Ala433	(1747)	HPPY-770-IRPY-VANE-VAR-BOOM-PO-700-	CTGTGGTACATCAAGATCTTC
Leu122-Ser199-Arg426-Lys432	(1771)		CTGTGGTACATCAAGATCTTC
Leu122-Ser199-Arg426-Gly431	(1771)		CIGIGGIACATCAAGAICTIC
Lys121-Val200-Asn425-Lys432	(1759)		CIGIGGIACATCAAGAICIIC
Val120-Ile201-Ile424-Ala433	(1747)	A CANONICA PROPERTY AND A CONTRACT OF THE PROPERTY OF THE PROP	CTGTGGTACATCAAGATCTTC
Val120-Ile201B-Ile424-Ala433	(1747)		CIGIGGTACATCAAGATCITC
Consensus	(1801)		CTGTGGTACATCAAGATCTTC
		1831	1860
Leu122-Ser199 Tryp427-Gly431	(1801)	ATCATGATC	GTGGGGGCCTGGTGGGGCCTG
Val127-Asn195-Arg426-Gly431	(1831)	ATCATGATC	GTGGGCGGCCTGGTGGGCCTG
Val120-Thr202-Ile424-Ala433	(1777)	ATCATGATO	Greececteteecote
Leu122-Ser199-Arg426-Lys432	(1801)		Greecectetteccete
Leu122-Ser199-Arg426-Gly431	(1801)		GTGGGCGCCTGGTGGGCCTG
Lys121-Val200-Asn425-Lys432	(1789)		GT GGGCGCCTGGTGGGCCTG
Val120-Ile201-Ile424-Ala433	(1777)	70.00	GTGGGCGGCCTGGTGGGCCTG
Val120-Ile201B-Ile424-Ala433	(1777)	BOX 2-17-12 - 17-2-2-2-2-2-2-1-1000-0-10-2-2-2-2-2-2-2-2	GTGGGCGGCCTGGTGGGCCTG
Consensus	(1831)	STATES AND	GTGGGCGGCCTGGTGGGCCTG
Consensus	(1001)	1861	1890
Leu122-Ser199 Tryp427-Gly431	(1831)		TTCACCGTGCTGAGCATCGTG
Vall27-Asn195-Arg426-Gly431			
	(1861)	COST CONTRACTOR CONTRACTOR CONTRACTOR CONTRACTOR	TTCACCGTGCTGAGCATCGTG
Val120-Thr202-Ile424-Ala433	(1807)		TTCACCGTGCTGAGCATCGTG
Leu122-Ser199-Arg426-Lys432	(1831)		TTCACCGTGCTGAGCATCGTG
Leu122-Ser199-Arg426-Gly431	(1831)	action in the second of the se	TTCACCGTGCTGAGCATCGTG
Lys121-Va1200-Asn425-Lys432	(1819)	CGCATCGTG	TTCACCGTGCTGAGCATCGTG

Vall20-Ile201-Ile424-Ala433	(1807)	CGCATCGTGTTCACCGTGCTGAGCATCGTG
Val120-Ile201B-Ile424-Ala433	(1807)	
Consensus	(1861)	
		1891 1920
Leu122-Ser199 Tryp427-Gly431	(1861)	1020
Val127-Asn195-Arg426-Gly431	(1891)	AACCGCGTGCGCCAGGGCTACAGCCCCCTG
Val120-Thr202-Ile424-Ala433	(1837)	AACCCCCMCCCCCAGGGC:ACAGCCCCCTG
Leu122-Ser199-Arg426-Lys432		
Louis Series Argazo Lysus 1	(1861)	
Leu122-Ser199-Arg426-Gly431	(1861)	
Lys121-Val200-Asn425-Lys432	(1849)	
Vall20-Ile201-Ile424-Ala433	(1837)	
Val120-Ile201B-Ile424-Ala433	(1837)	
Consensus	(1891)	AACCGCGTGCGCCAGGGCTACAGCCCCCTG
		1921 1950
Leu122-Ser199 Tryp427-Gly431	(1891)	AGCTTCCAGACCCGCTTCCCGGCCCCCGC
Val127-Asn195-Arg426-Gly431	(1921)	ACCUMPACA CARCAGO CONTRACTOR CONT
Val120-Thr202-Ile424-Ala433	(1867)	AGCTTCCAGACCCGCTTCCCCGCCCCCGC
Leu122-Ser199-Arg426-Lys432		AGCTTCCAGACCCGCTTCCCCGCCCCCGC
Leu122-Sei 199-Arg426-Lys432	(1891)	AGCTTCCAGACCCGCTTCCCCGCCCCCGC
Leu122-Ser199-Arg426-Gly431	(1891)	AGCITCCAGACCCGCTTCCCCGCCCCCCGC
Lys121-Val200-Asn425-Lys432	(1879)	AGCTTCCAGACCCGCTTCCCCGCCCCCGC
Val120-Ile201-Ile424-Ala433	(1867)	AGCTTCCAGACCCGCTTCCGCGCCCCCGC
Val120-Ile201B-Ile424-Ala433	(1867)	AGCTTCCAGACCCGCTTCCCCGCCCCCGC
Consensus	(1921)	AGCTTCCAGACCCGCTTCCCCGCCCCCCGC
		1951 1980
Leu122-Ser199 Tryp427-Gly431	(1921)	GGCCCGACCGCCCGAGGGCATCGAGGAG
Val127-Asn195-Arg426-Gly431	(1951)	GGCCCCACCGCCCCGAGGGCATCGAGGAG
Val120-Thr202-Ile424-Ala433	(1897)	
Leu122-Ser199-Arg426-Lys432		GGCCCCGACCGCCCCGAGGGCATCGAGGAG
Low122 Car100 Arrivator 23 421	(1921)	GGCCCGACCGGCCCGAGGCCATCGAGGAG
Leu122-Ser199-Arg426-Gly431	(1921)	GGCCCGACCGCCCCGAGGGCATCGAGGAG
Lys121-Val200-Asn425-Lys432	(1909)	Geolegica Accession con a Cector And Control
Val120-Ile201-Ile424-Ala433	(1897)	GGEC CEL CEU COMPACEMENTEMA
Val120-Ile201B-Ile424-Ala433	(1897)	GGCCCCGACCGC CCCACCCCATCGACGAG
Consensus	(1951)	GGCCCCGACCGCCCCGAGGGCATCGAGGAG
		1981 2010
Leul22-Ser199 Tryp427-Gly431	(1951)	GAGGGCGCCAGGGCAGGGCAGGGCAGG
Val127-Asn195-Arg426-Gly431	(1981)	GAGGGCGGGAGCGGAGCGCAGC
Val120-Thr202-Ile424-Ala433	(1927)	GAGGGGGGGAGCGCAGCGCGAGCGCAGC
Leu122-Ser199-Arg426-Lys432	(1951)	CASCISCIONALE COLORA CO
Leu122-Ser199-Arg426-Gly431		GAGGGGGGGAGGGGAGCGAGC
	(1951)	GAGGGGGGGAGGGGAGGGAGG
Lys121-Val200-Asn425-Lys432	(1939)	GAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
Val120-Ile201-Ile424-Ala433	(1927)	GAGGGGGGCGAGGCAGCGGAGCGCAGC
Vall20-Ile201B-Ile424-Ala433	(1927)	GAGGECGGCGAGCGGACCGCAGC
Consensus	(1981)	GAGGGCGAGCGCGACCGCAGC
		2011 2040
Leu122-Ser199 Tryp427-Gly431	(1981)	AGCCCCTGGTGCACGGCGTGGTGGCCTG
Val127-Asn195-Arg426-Gly431	(2011)	AGCCCCCTGCTEGACGGCCTCGTGCCCTG
Val120-Thr202-Ile424-Ala433	(1957)	AGCCCCTGGTGCACGGCCTGCTGCCCTG
Leu122-Ser199-Arg426-Lys432	(1981)	ACCOCCUSS GOALGGE GET GET CET G
Leu122-Ser199-Arg426-Gly431		AGCCCCTGGTGCACGGCCTGCTGGCCCTG
Lys121-Val200-Asn425-Lys432	(1981)	AGCCCCTEGIGCACGGCCTGCTGGCCCTG
Upli20 Tl-201 Tr 404 Tr	(1969)	AGCCCCTGGTGC#CGGCCTGCTGGCCCTG
Val120-Ile201-Ile424-Ala433	(1957)	AGCCCCTGGTGCACGGCCTGCTGGCCCTG
Val120-Ile201B-Ile424-Ala433	(1957)	AGCCCCTGGTGCACGGCCTGCTGGCCCTG
Consensus	(2011)	AGCCCCTGGTGCACGGCCTGCTGGCCCTG
		2041 2070
Leu122-Ser199 Tryp427-Gly431	(2011)	ATCTGGGAGGAGCTGTGGCTG
Val127-Asn195-Arg426-Gly431	(2041)	ATCTGGGACGACCTGCGCAGCCTGTGCCTG
Val120-Thr202-Ile424-Ala433	(1987)	ATCTGGGAGGAGCTGCGCAGCCTGTGCCTG
	, /	G TO THE STATE OF

Leu122-Ser199-Arg426-Lys432	(2011)	ATCTGGGACGACCTGCGCAGCCTGTGCCTG
Leu122-Ser199-Arg426-Gly431	(2011)	ATCTGGGACGACCTGCGCAGCCTGTGCCTG
Lys121-Val200-Asn425-Lys432	(1999)	ATCTGGGACGACCTGCGCAGCCTGTGCCTG
Vall20-Ile201-Ile424-Ala433	(1987)	ATCTGGGACGACCTGCGCAGCCTGTGCCTG
Vall20-Ile201B-Ile424-Ala433	(1987)	ATCTGGGACGACCTGCGCAGCCTGTGCCTG
Consensus	(2041)	ATCTGGGACGACCTGCGCAGCCTGTGCCTG
		2071 2100
Leul22-Ser199 Tryp427-Gly431	(2041)	TTCAGCTACCACCGCCTGCGCGACCTGATC
Val127-Asn195-Arg426-Gly431	(2071)	TTCAGCTACCACCGCCTGCGCGACCTGATC
Val120-Thr202-Ile424-Ala433	(2017)	TTCAGCTACCACCGCCTGCGCGACCTGATC
Leu122-Ser199-Arg426-Lys432	(2041)	TTCAGCTACCACCGCCTGCGCGACCTGATC
Leu122-Ser199-Arg426-Gly431	(2041)	TTCAGCTACCACCGCCTGCGCGACCTGATC
Lys121-Val200-Asn425-Lys432	(2029)	TTCAGCTACCACCGCCTGCGCGACCTGATC
Val120-Ile201-Ile424-Ala433	(2017)	TTCAGCTACCACCGCCTGCGCGACCTGATC
Val120-Ile201B-Ile424-Ala433	(2017)	TTCAGCTACCACCGCCTGCGCGACCTGATC
Consensus	(2071)	TTCAGCTACCACCGCCTGCGCGACCTGATC
		2101 2130
Leu122-Ser199 Tryp427-Gly431	(2071)	CTGATCGCCGCCCGCATCGTGGAGCTGCTG
Val127-Asn195-Arg426-Gly431	(2101)	CTGATCGCCGCCCGCATCGTGGAGCTGCTG
Val120-Thr202-Ile424-Ala433	(2047)	CTGATCGCCGCCCGCATCGTGGAGCTGCTG
Leu122-Ser199-Arg426-Lys432	(2071)	CTGATCGCCCCCCCCATCGTGGAGCTGCTG
Leu122-Ser199-Arg426-Gly431	(2071)	CTGATCGCCGCCCGCATCGTGGAGCTGCTG
Lys121-Val200-Asn425-Lys432	(2059)	CTGATCGCCGCCCGCATCGTGGAGCTGCTG
Val120-Ile201-Ile424-Ala433	(2047)	CIGATOGCCGCCGCATCGTGGAGCTGCTG
Val120-Ile201B-Ile424-Ala433	(2047)	CIGATOGCOGCCCGCATCGTGGAGCTGCTG
Consensus	(2101)	CTGATCGCCGCCCGCATCGTGGAGCTGCTG
		2131 2160
Leu122-Ser199 Tryp427-Gly431	(2101)	GGCCGCCGCGGCTGGGAGGCCCTGAAGTAC
Val127-Asn195-Arg426-Gly431	(2131)	GGCCGCCGCGGCTGGGGGCCCTGAAGTAC
Val120-Thr202-Ile424-Ala433	(2077)	GGCCGCCGCGGCTGGGAGCCCCTGAAGTAC
Leu122-Ser199-Arg426-Lys432	(2101)	GGCCGCCGGGCTGGGAGGCCCCTGAAGTAC
Leu122-Ser199-Arg426-Gly431	(2101)	GGCCGCCGCGCTGGGAGGCCCTGAAGTAC
Lys121-Val200-Asn425-Lys432	(2089)	GGGGGCCCGGGCTGGGGCCCCTGAAGTAC
Vall20-Ile201-Ile424-Ala433	(2077)	GGCCGCCGGGCTGGGAGCCCCTGAAGTAC
Vall20-Ile201B-Ile424-Ala433	(2077)	GGCCGCCGCGCTGGAGTAC
Consensus	(2131)	GGCCGCCGCGCTGGGGGCCCTGAAGTAC
		2161 2190
Leu122-Ser199 Tryp427-Gly431	(2131)	TGGGGCAACCTGCTGCAGTACTGGATCCAG
Val127-Asn195-Arg426-Gly431	(2161)	TGGGGCAACCTGCTGCAGTACTGGATCCAG
Val120-Thr202-Ile424-Ala433	(2107)	TGGGGCARCCTGCTGCAGTACTGGATCCAG
Leu122-Ser199-Arg426-Lys432	(2131)	TGGGGCAACGEGCTGCAGTACTGGATCCAG
Leu122-Ser199-Arg426-Gly431	(2131)	TGGGGCAACCTGCTGCAGTACTGGATCCAG
Lys121-Val200-Asn425-Lys432	(2119)	TGGGGCAAGCTGCTGCAGTACTGGATCCAG
Val120-Ile201-Ile424-Ala433	(2107)	TGGGGGAACCIGCTGCAGTACTGGATCCAG
Val120-Ile201B-Ile424-Ala433	(2107)	TGGGGGAACCTGCTGCAGTACTGGATECAG
Consensus	(2161)	TGGGGCAACCTGCTGCAGTACTGGATCCAG
		2191 2220
Leu122-Ser199 Tryp427-Gly431		GAGCTGAAGAACAGCGCCGTGAGCCTGTTC
Val127-Asn195-Arg426-Gly431		GAGCTGAAGAACAGCGCCGTGAGCCTGTTC
Val120-Thr202-Ile424-Ala433	(2137)	GAGCTGAAGAACAGCGCCGTGAGCCTGTTC
Leu122-Ser199-Arg426-Lys432		GAGCTGAAGAACAGCGCCGTGAGCCTGTTC
Leu122-Ser199-Arg426-Gly431		GAGCTGAAGAACAGCGCCGTGAGCCTGTTC
Lys121-Val200-Asn425-Lys432		GAGCTGAAGAACAGCGCCGTGAGCCTGTTC
Val120-Ile201-Ile424-Ala433		GAGCTGAAGAACAGCGCCGTGAGCCTGTTC
Vall20-Ile201B-Ile424-Ala433		GAGCTGAAGAACAGCGCCGTGAGCCTGTTC
Consensus	(2191)	GAGCTGAAGAACAGCGCCGTGAGCCTGTTC
		2221 2250

WO 00/39303	41	/ 65 PCT/US99/312	72
Leu122-Ser199 Tryp427-Gly431	(2191)) GACGCCATCGCCATCGCCGTGGCCGAGGGC	
Val127-Asn195-Arg426-Gly431	(2221)		
Val120-Thr202-Ile424-Ala433	(2167)		
Leu122-Ser199-Arg426-Lys432	(2191)		
Leu122-Ser199-Arg426-Gly431	(2191)		
Lys121-Val200-Asn425-Lys432	(2179)		
Val120-Ile201-Ile424-Ala433	(2167)		
Val120-Ile201B-Ile424-Ala433	(2167)		
Consensus	(2221)		
		2251 2280	
Leul22-Ser199 Tryp427-Gly431	(2221)) ACCGACCGCATCATCGAGGTGGCCCAGCGC	
Val127-Asn195-Arg426-Gly431	(2251)		
Vall20-Thr202-Ile424-Ala433	(2197)) ACCGACCGCATCATCGAGGTGGCCCAGCGC	
Leu122-Ser199-Arg426-Lys432	(2221)) ACCGACCGCATCATCGAGGTGGCCCAGCGC	
Leu122-Ser199-Arg426-Gly431	(2221)		
Lys121-Val200-Asn425-Lys432	(2209)) ACCGACCGCATCATCGAGGTGGCCCAGCGC	
Val120-Ile201-Ile424-Ala433	(2197)) ACCGACCGCATCATCGAGGTGGCCCAGCGC	
Vall20-Ile201B-Ile424-Ala433	(2197)		
Consensus	(2251)) ACCGACCGCATCATCGAGGTGGCCCAGCGC	
		2281 2310	
Leu122-Ser199 Tryp427-Gly431	(2251)	Control of the Contro	
Val127-Asn195-Arg426-Gly431	(2281)	**メルフルがイ本があることものがあることをあることをおかりのない。ことは、ことは、ことは、ことは、ことは、ことは、ことは、ことは、ことは、ことは、	
Val120-Thr202-Ile424-Ala433	(2227)		
Leu122-Ser199-Arg426-Lys432	(2251)		
Leu122-Ser199-Arg426-Gly431	(2251)	0.000 0.000	
Lys121-Val200-Asn425-Lys432	(2239)		
Vall20-Ile201-Ile424-Ala433	(2227)		
Val120-Ile201B-Ile424-Ala433	(2227)		
Consensus	(2281)		
		2311 2340	
Leul22-Ser199 Tryp427-Gly431	(2281)		
Val127-Asn195-Arg426-Gly431	(2311)	Section Control Contro	
Val120-Thr202-Ile424-Ala433	(2257)		
Leu122-Ser199-Arg426-Lys432	(2281)		
Leu122-Ser199-Arg426-Gly431	(2281)	**************************************	
Lys121-Val200-Asn425-Lys432	(2269)	A STATE OF THE PROPERTY OF THE	
Val120-Ile201-Ile424-Ala433	(2257)		
Val120-Ile201B-Ile424-Ala433	(2257)		
Consen s us	(2311)		
T122 G100 M407 G3421		2341 2352	
Leu122-Ser199 Tryp427-Gly431	(2311)		
Vall27-Asn195-Arg426-Gly431	(2341)		
Vall20-Thr202-Ile424-Ala433	(2287)		
Leu122-Ser199-Arg426-Lys432 Leu122-Ser199-Arg426-Gly431	(2311)		
Lys121-Val200-Asn425-Lys432	(2311)		
Val120-Ile201-Ile424-Ala433	(2299) (2287)		
Val120-Ile2018-Ile424-Ala433			
Consensus	(2287) (2341)	min a committee and the state of the state o	
Consensus	(2041)	CIGIAACICOAG	

SEQ ID NO:3 VAL120-ALA204

GAATTCGCCACCATGGATGCAATGAAGAGGGGCTCTGCTGTGTGCTGCTGTTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGGGCGCCGGCGCCTGCCCCAA GGTGAGCTTCGAGCCCATCCCATCCACTACTGCGCCCCGCCGGCTTCGCCATCCTGAAGTG CAACGACAAGAAGTTCAACGCCAGCGCCCCTGCACCAACGTGAGCACCGTGCAGTGCACCC ACGGCATCCGCCCGTGGTGAGCACCCAGCTGCTGAACGGCAGCCTGGCCGAGGAGGGC GTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGA GAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGAGCATCACCATCGGCC CCGGCCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACA TCAGCGGCGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTC GGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAG CTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAA CAACACCATCGGCCCAACAACACCAACGGCACCATCACCCTGCCCTGCCGCATCAAGCAGA TCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATGTACGCCCCCCCATCCGCGGCCAGATC CGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGAAGGAGATCAGCAA CACCACGAGATCTTCCGCCCGGCGGCGGCGACATGCGCGACAACTGGCGCAGCGAGCTGT ACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAAGGCCAAGCGCCGC GTGGTGCAGCGCGAGAAGCGCGCCGTGACCCTGGGCGCCATGTTCCTGGGCTTCCTGGGCGCC GCCGCCACCATGGCCCCGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGAG AGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTG AAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGT GCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAACATGACCTGGA TGGAGTGGGAGCGCGAGATCGACAACTACACCAACCTGATCTACACCCTGATCGAGGAGAGC CAGAACCAGCAGGAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGT GGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCG GCCTGGTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCT ACAGCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGCGCCCCGACCGCCCCGAGGGCA TCGAGGAGGAGGCGCGACCGCACCGCACCCCCTGGTGCACGCCTGCTG ${\tt GCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCCTGCGCGACCTG}$ ATCCTGATCGCCGCCGCATCGTGGAGCTGCTGGGCCGCCGCGGGTGGGAGGCCCTGAAGTAC TGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGA $\tt CGCCATCGCCATCGCCGAGGGCACCGACCGCATCATCGAGGTGGCCCAGCGCATCG$ GCCGCGCCTTCCTGCACATCCCCGCCGCATCCGCCAGGGCTTCGAGCGCCCTGCTGTAAC **TCGAG**

SEO ID NO:4 VAL120-ILE201

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGGGCGGCATCACCCAGGCCTG CCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGCCCCGGCGGCTTCGCCATCCT GAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGTGAGCACCGTGCAGT GCACCCACGCCTCGTCGTGAGCACCCAGCTGCTGAACGGCAGCCTGGCCGAG GAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCT GAAGGAGAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGAGCATCACCA TCGGCCCGGCCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACT GCAACATCAGCGGCGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCC CAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGAT GCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCAC CTGGAACACCATCGGCCCCAACACACCAACGCACCATCACCCTGCCCTGCCGCATCA AGCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATGTACGCCCCCCCATCCGCGGC CAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGAGAT CAGCAACACCACCGAGATCTTCCGCCCCGGCGCGGCGACATGCGCGACAACTGGCGCAGCG AGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCACCAAGGCCAAG CGCCGCGTGGTGCAGCGCGAGAAGCGCGCCGTGACCCTGGGCGCCATGTTCCTGGGCTTCCTG GCTGAGCGGCATCGTGCAGCAGCAGAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACC TGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGC TACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCAC CGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAACATGA CCTGGATGGAGTGGGAGCGCGAGATCGACAACTACACCAACCTGATCTACACCCTGATCGAG GAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCA GCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCG TGGGCGGCTGGTGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCC AGGGCTACAGCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGCGGCCCCGACCGCCCCG AGGGCATCGAGGAGGAGGGCGCGAGCGCGACCGCACCAGCAGCCCCCTGGTGCACGG CCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCCTGCG CGACCTGATCCTGATCGCCGCCCGCATCGTGGAGCTGCTGGGCCGCCGCGCGGCTGGGAGGCCCT GAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCC TGTTCGACGCCATCGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCCAGC GCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCCCTGC **TGTAACTCGAG**

SEQ ID NO:5 VAL120-ILE201B

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCAGTCTTCG TTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCA CCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGTGCACAACGTGTGGGCCACCC ACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACA TGTGGAAGAACAACATGGTGGAGCAGATGCACGAGGACATCATCAGCCTGTGGGACCAGAGCCTGAAGC CCTGCGTGCCCGGCATCACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGC CCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGT GAGCACCGTGCAGTGCACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCT GGCCGAGGGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCT GAAGGAGAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGAGCATCACCATCGGCCC CGGCCGCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGC GAGAAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGACCATC GTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCGGCGGCGAGTTC TTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCGGCCCCAACAACACCAAC GGCACCATCACCCTGCCCTGCCGCATCAAGCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG TACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACG GCGGCAAGGAGATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGACAACTGGC GCAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAAGGCCAAGC GCCGCGTGGTGCAGCGCGAGAAGCGCGCCGTGACCCTGGGCCCATGTTCCTGGGCTTCCTGGGCGCCGC CGGCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGAGCGGCATCGT GCAGCAGCAGAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGG CATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCAT ${\tt CTGGGGCTGCAGCGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAG}$ CCTGGACCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACTACACCAACCT GATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGG ACAAGTGGGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCAT GATCGTGGGCCTGGTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAG GGCTACAGCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCCGGGCCCCGACCGCCCCGAGGGCATCG AGGAGGAGGGCGAGCGCGACCGCACCGCAGCAGCCCCCTGGTGCACGGCCTGCTGGCCCTGATCT GGGACGACCTGCGCAGCCTGTTCAGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCG CATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGCAGTACTG GATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCGCCGTGGCCGAGGGCAC CGACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCCCTTCCTGCACATCCCCCGCCGCATCCGCCAG GGCTTCGAGCGCCCCTGCTGTAACTCGAGCGTGCT

SEQ ID NO:6 LYS121-VAL200

GAATTCGCCACCATGGATGCAATGAAGAGGGGCTCTGCTGTTGTGCTGCTGTTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGCCACCACCCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGGCCCCCGTGATCACCCA GGCTGCCCAAGGTGAGCTTCGAGCCCATCCACTACTGCGCCCCGCCGGCTTCGC CATCCTGAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGTGAGCACCG TGCAGTGCACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGCAACGGCAGCCTGG CCGAGGAGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTG CAGCTGAAGGAGAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGAGCAT CACCATCGGCCCGGCCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGC CCACTGCAACATCAGCGGCGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGC AGGCCCAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATC GTGATGCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAAC AGCACCTGGAACAACACCATCGGCCCCAACAACACCAACGGCACCATCACCCTGCCCTGCCG CATCAAGCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATGTACGCCCCCCCATCC GCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAG GAGATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGACAACTGGCG CAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCACCAAGG CCAAGCGCCGCGTGGTGCAGCGCGAGAAGCGCGCCGTGACCCTGGGCGCCATGTTCCTGGGC TTCCTGGGCGCCGCCGGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCAGGCCCGC CAGCTGCTGAGCGCATCGTGCAGCAGCAGAACAACCTGCTGCGCGCCATCGAGGCCCAGCA GCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGG AGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGC ACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAA CATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACTACACCAACCTGATCTACACCCTGA TCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGACCAAGTG GGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCAT GATCGTGGGCGGCCTGGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGT GCGCCAGGGCTACAGCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCCGCGCCCCGACCG CCCCGAGGGCATCGAGGAGGAGGGCGCGAGCGCGACCGCAGCAGCAGCCCCCTGGTGC ACGGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCC TGCGCGACCTGATCCTGATCGCCGCCCGCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGG CCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTG CAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCC CTGCTGTAACTCGAGCGTGCT

SEQ ID NO:7: LEU122-SER199

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA $\tt GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG$ TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCTGCGTGCCCACCGACCCCAACCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGGGCAACAGCGTGAT CTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCAGCGCCCCTGCACCAACGTGA GCACCGTGCAGTGCACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGCAACGGC AGCCTGGCCGÁGGAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCAT CATCGTGCAGCTGAAGGAGAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCA AGAGCATCACCATCGGCCCGGCCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCC GCCAGGCCCACTGCAACATCAGCGGCGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACC AAGCTGCAGGCCCAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCC CGAGATCGTGATGCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCT GTTCAACAGCACCTGGAACAACACCATCGGCCCCAACAACACCAACGGCACCATCACCCTGC CCTGCCGCATCAAGCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATGTACGCCCCC GGCAAGGAGATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGACAA $\tt CTGGCGCAGCGAGCTGTACAAGTACAAGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCA$ ${\tt CCAAGGCCAAGCGCGCGTGTGCAGCGCGAGAAGCGCGCGTGACCCTGGGCGCATGTTC}$ GCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCAGCACCACCTGCTGCGCCCCATCGAGGC ATCTGCACCACCGCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTG GAACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACTACACCAACCTGATCTACA CCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGA CAAGTGGGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTT CATCATGATCGTGGGCCTGGTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAA CCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCCGCGGCCC CGACCGCCCGAGGGCATCGAGGAGGAGGGCGCGACCGCGACCGCAGCAGCAGCCCC ${\tt CTGGTGCACGGCCTGGTCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGCTAC}$ CACCGCTGCGCGACCTGATCCTGATCGCCGCCCGCATCGTGGAGCTGCTGGGCCGCCGCGGC TGGGAGCCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAG GGTGGCCCAGCGCATCGCCGCCTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGA GCGCGCCCTGCTGTAACTCGAGCGTGCT

SEQ ID NO:8 VAL120-THR202

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGGGCGGCGCCACCCAGGCCTG CCCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTGCGCCCCGCCGGCTTCGCCATCCT GAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGTGAGCACCGTGCAGT GCACCCACGGCATCCGCCCGTGGTGAGCACCCAGCTGCTGAACGCAGCCTGGCCGAG GAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCT GAAGGAGAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGAGCATCACCA TCGGCCCGGCCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACT GCAACATCAGCGGCGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCC CAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGAT GCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCAC CTGGAACACACCATCGGCCCCAACAACACCAACGGCACCATCACCCTGCCCTGCCGCATCA CAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGAGAT CAGCAACACCACGAGATCTTCCGCCCGGCGGCGGCGACATGCGCGACAACTGGCGCAGCG AGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCACCAAGGCCAAG CGCCGCGTGGTGCAGCGCGAGAAGCGCGCCGTGACCCTGGGCGCCATGTTCCTGGGCTTCCTG GGCGCCGCCGCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCT GCTGAGCGCATCGTGCAGCAGCAGCACCACCTGCTGCGCGCCATCGAGGCCCAGCAGCACC TGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGC TACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCAC CGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAACATGA CCTGGATGGAGTGGAGCGCGAGATCGACAACTACACCAACCTGATCTACACCCTGATCGAG GAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCA GCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCG TGGGCGGCCTGGTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCC AGGCTACAGCCCCTGAGCTTCCAGACCGCTTCCCCGCCCCCGCGGCCCCGACCGCCCCG AGGGCATCGAGGAGGGGGGGGGGGGGGGCGACCGCAGCAGCAGCCCCCTGGTGCACGG CCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCCTGCG CGACCTGATCCTGATCGCCGCCCGCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCT GAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCC TGTTCGACGCCATCGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCCAGC GCATCGGCCGCCCTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCCCTGC TGTAACTCGAG

SEO ID NO:9 TRP427-GLY431

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG ACCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGA GCCCATCCCATCCACTACTGCGCCCCGCCGCTTCGCCATCCTGAAGTGCAACGACAAGAA GTTCAACGCCAGCGCCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCC CCGTGGTGAGCACCCAGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC AGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGCGTGGAGAT CAACTGCACCGCCCAACAACACCCGCAAGAGCATCACCATCGGCCCCGGCCGCCCT TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG AAGTGGAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGAC CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG GCCCCAACACCCAACGGCACCATCACCCTGCCCTGCCGCATCAAGCAGATCATCAACCGCT GGGGCGCAAGGCCATGTACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACATC ACCGGCTGCTGCTGACCCGCGACGGCGCAAGGAGATCAGCAACACCACCGAGATCTTCCG CCCCGCGCGCGCCACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGTGGTGA AGATCGAGCCCTGGGCGTGGCCCCCACCAAGGCCAAGCGCCGCGTGGTGCAGCGCGAGAAG CGCGCCGTGACCCTGGGCGCCATGTTCCTGGGCTTCCTGGGCGCCGCCGGCAGCACCATGGGC GCCCGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCA GAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCA TCAAGCAGCTGCAGGCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTG ATCGACAACTACACCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAA GAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGACATCA GCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGCA TCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCC AGACCCGCTTCCCCGCCCCCCGCGGCCCCGACCCCCCGAGGGCATCGAGGAGGGGGGGCGGC GAGCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACGA CCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCG CATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGC AGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCGCC GTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCCCTTCCTGCA CATCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG

SEQ ID NO:10 ARG426-GLY431

GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCTGTGCGTG ACCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGA GCCCATCCCATCCACTACTGCGCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAA GTTCAACGGCAGCGGCCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCC CCGTGGTGAGCACCCAGCTGCTGCAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC AGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT CAACTGCACCCGCCCAACAACAACACCCGCAAGAGCATCACCATCGGCCCGGCCGCCCT TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG AAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGAC CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG GCCCCAACACCCAACGGCACCATCACCCTGCCCTGCCGCATCAAGCAGATCATCAACCGC GGCGGCGCAAGGCCATGTACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACAT CACCGGCCTGCTGACCCGCGACGGCGCAAGGAGATCAGCAACACCACCGAGATCTTCC GCCCGGCGGCGACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGTGGTG AAGATCGAGCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGCGTGGTGCAGCGCGAGAA GCGCGCGTGACCCTGGGCGCATGTTCCTGGGCTTCCTGGGCGCCGCCGGCAGCACCATGGG CGCCGCAGCCTGACCTGACCGTGCAGGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGC AGAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGC ATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCT GGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCT GATCGACAACTACACCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAGA AGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGACATC AGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGC ATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTC CAGACCGCTTCCCGCCCCCGCGGCCCCGACCGCCCCGAGGGCATCGAGGAGGAGGGGGG CGAGCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACG ACCTGCGCAGCCTGTTCCAGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCC GCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGAGGCCCTGAAGTACTGGGGCAACCTGCTG CAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCGC CGTGGCCGAGGGCACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCGCTTCCTGC ACATCCCCGCCGCATCCGCCAGGGCTTCGAGCGCCCCTGCTGTAACTCGAG

SEQ ID NO:11 ARG426-GLY431B

GAATTCGCCACCATGGATGCAATGAAGAGGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGCCACCACCCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG ACCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT GGACCGCGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGCCCCATCGACAACGACAACACCAGC TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGA GCCCATCCCATCCACTACTGCGCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAA GTTCAACGCAGCGCCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCC CCGTGGTGAGCACCCAGCTGCTGCACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC AGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT CAACTGCACCGCCCAACAACAACACCCGCAAGAGCATCACCATCGGCCCCGGCCGCCCT TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG AAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGAC CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG GCCCCAACACCCACGGCACCATCACCCTGCCCTGCCGCATCAAGCAGATCATCAACCGC GGCAGCGGCAAGGCCATGTACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACAT CACCGGCCTGCTGCTGACCCGCGACGGCGACAGGAGATCAGCAACACCACCGAGATCTTCC GCCCGGCGCGCGACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGTGGTG AAGATCGAGCCCCTGGGCGTGGCCCCCACCAAGGCCAAGCGCCGCGTGGTGCAGCGCGAGAA GCGCGCGTGACCCTGGGCGCATGTTCCTGGGCTTCCTGGGCGCCGCCGGCAGCACCATGGG CGCCGCAGCCTGACCCTGCCGGCCAGCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGC AGAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGC ATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCT GGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCT GATCGACAACTACACCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAGA AGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGACATC AGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGC ATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCCAGGGCTACAGCCCCCTGAGCTTC CAGACCGCTTCCCGCCCCCCGCGCCCCGACCGCCCCGAGGGCATCGAGGAGGAGGGGGG CGAGCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACG ACCTGCGCAGCCTGTTCCAGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCC GCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTG CAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCGC CGTGGCCGAGGGCACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCCTTCCTGC ACATCCCCGCCGCATCCGCCAGGGCTTCGAGCGCCCCTGCTGTAACTCGAG

SEQ ID NO:12 ARG426-LYS432

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG ACCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGA GCCCATCCCATCCACTACTGCGCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAA GTTCAACGCAGCGCCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCC ${\tt CCGTGGTGAGCACCCAGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC}$ AGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT CAACTGCACCGCCCAACAACACCCGCAAGAGCATCACCATCGGCCCGGCCGCCCT TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG AAGTGGAACAACACCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGAC CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG GCCCCAACACACCAACGGCACCATCACCCTGCCCTGCCGCATCAAGCAGATCATCAACCGC GGCGGCAACAAGGCCATGTACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACAT CACCGCCTGCTGCTGACCCGCGACGCCGCAAGGAGATCAGCAACACCACCGAGATCTTCC GCCCGGCGGCGACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGTGGTG AAGATCGAGCCCTGGGCGTGGCCCCCACCAAGGCCAAGCGCGCGTGGTGCAGCGCGAGAA GCGCGCGTGACCCTGGGCGCATGTTCCTGGGCTTCCTGGGCGCCGCCGCAGCACCATGGG CGCCGCAGCCTGACCCTGACGCTGCAGGCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGC AGAACAACCTGCTGCGCCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGC ATCAAGCAGCTGCAGGCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCT GGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCT GATCGACAACTACACCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAGA AGAACGAGCAGGAGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGACATC AGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGC ATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTC CAGACCGCTTCCCGCCCCCGCGGCCCCGACGCCCCGAGGGCATCGAGGAGGAGGGCGG CGAGCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACG ACCTGCGCAGCCTGTTCCGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCC GCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTG CAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCGC CGTGGCCGAGGGCACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCCTTCCTGC ACATCCCCGCCGCATCCGCCAGGGCTTCGAGCGCCCCTGCTGTAACTCGAG

SEQ ID NO:13 ASN425-LYS432

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG ACCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGA GCCCATCCCCATCCACTACTGCGCCCCGCCGCCTTCGCCATCCTGAAGTGCAACGACAAGAA GTTCAACGCCAGCGCCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCC CCGTGGTGAGCACCCAGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC AGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT CAACTGCACCGCCCCAACACACACCCCGCAGGGCATCACCATCGGCCCCGGCCGCCCC TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG **AAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGAC** CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG GCCCAACAACACCAACGGCACCATCACCCTGCCCTGCCGCATCAAGCAGATCATCAACGCCC CCAAGGCCATGTACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCC TGCTGCTGACCCGCGACGGCGCAAGGAGATCAGCAACACCACCGAGATCTTCCGCCCCGGC GGCGGCGACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGA GCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGCGTGGTGCAGCGCGAGAAGCGCGCG TGACCCTGGGCGCATGTTCCTGGGCTTCCTGGGCGCCGCCGGCAGCACCATGGGCGCCCGCA GCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCAGCAGCACAAC CTGCTGCGCCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCA GCTGCAGGCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCT GGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAAC CTACACCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGC AGGAGCTGCTGGACCAGTGGGCCAGCCTGTGGACCTGGTTCGACATCAGCAAGTGG CTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGCATCGTGTTC ACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGC CCGCGACCGCAGCACCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAG GCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGA TCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCGCCGTGGCCGAG GGCACCGACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGC CGCATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG

SEQ ID NO:14 ILE424-ALA433

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTTGTGCTGCTGTTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG ACCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGA GCCCATCCCATCCACTACTGCGCCCCGCCGCTTCGCCATCCTGAAGTGCAACGACAAGAA GTTCAACGGCAGCGCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCC CCGTGGTGAGCACCCAGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC AGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT CAACTGCACCCGCCCAACAACAACACCCGCAAGAGCATCACCATCGGCCCCGGCCGCCCT TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG AAGTGGAACAACACCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGAC CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG GCCCCAACACCCACGGCACCATCACCCTGCCCTGCCGCATCAAGCAGATCATCGGCGGC GCCATGTACGCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTG CTGACCGGGGGGGGGAAGGAGATCAGCAACACCACCGAGATCTTCCGCCCGGCGGCGG CGACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGTGGTGAAGATCGAGCCCC TGGGCGTGGCCCCACCAAGGCCAAGCGCCGCGTGGTGCAGCGCGAGAAGCGCGCGTGACC CTGGGCGCCATGTTCCTGGGCTTCCTGGGCGCCGCCGCAGCACCATGGGCGCCCGCAGCCTG ACCCTGACCGTGCAGGCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCAGCAGCACCTGCT GCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCCATCAAGCAGCTGC AGGCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGC TGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAG CCTGGACCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACTACA CCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGA GCTGCTGGACCAGTGGGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGT GGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGCATCGTGTTCACCG TGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCC GACCGCAGCAGCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTG CTGGGCCGCGGGTGGGAGGCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCA GGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCGCCGTGGCCGAGGGCA CCGACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCCTTCCTGCACATCCCCCGCCGCA TCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG

SEQ ID NO:15 ILE423-MET434

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG ACCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGA GCCCATCCCATCCACTACTGCGCCCCGCCGCTTCGCCATCCTGAAGTGCAACGACAAGAA GTTCAACGCAGCGCCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCC CCGTGGTGAGCACCCAGCTGCTGAACGCAGCCTGCCGAGGAGGGCGTGGTGATCCGC AGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT CAACTGCACCGCCCAACAACACCCGCAAGAGCATCACCATCGGCCCGGCCGCCCT TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG AAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGAC ${\tt CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG}$ GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG TACGCCCCCCATCCGCGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGACC CGCGACGGCGGCAAGGAGATCAGCAACACCACGAGATCTTCCGCCCCGGCGGCGGCGACAT GCGCGACAACTGGCGCGGCGGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCG TGGCCCCACCAAGGCCAAGCGCCGCGTGGTGCAGCGCGAGAAGCGCGCGTGACCCTGGGC GCCATGTTCCTGGGCTTCCTGGGCGCCGCCGGCACCCTGGGCGCCCGCAGCCTGACCCTG ACCGTGCAGGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCAGAACAACCTGCTGCGCGC CATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCC GCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGC GGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGA CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACTACACCAACC TGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTG GAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACAT CAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGCATCGTGTTCACCGTGCTGAG CATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCC AGCAGCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTG CGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCT GAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCGCCGTGGCCGAGGGCACCGACC GCATCATCGAGGTGGCCCAGCGCATCGGCCGCCCTTCCTGCACATCCCCCGCCGCATCCGCC AGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG

SEQ ID NO:16 GLN422-TYR435

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCTGCGTGAAGCTGACCCCCCTGTGCGTG ACCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGGAT GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGA GCCCATCCCATCCACTACTGCGCCCCGCCGCTTCGCCATCCTGAAGTGCAACGACAAGAA GTTCAACGCCAGCGCCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCC CCGTGGTGAGCACCCAGCTGCTGAACGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC AGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT CAACTGCACCGCCCAACAACACCCGCAAGAGCATCACCATCGGCCCGGCCGCCCT TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG AAGTGGAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGAC CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG GCCCAACAACACCAACGCACCATCACCTGCCCTGCCGCATCAAGCAGGGCGGCTACGCC CCCCCATCCGCGCCAGATCCGCTGCAGCAGCACATCACCGGCCTGCTGACCCGCGAC GGCGGCAAGGAGATCAGCAACACCACGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGA CAACTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCC CCACCAAGGCCAAGCGCGCGTGGTGCAGCGCGAGAAGCGCGCGTGACCCTGGGCGCCATG TTCCTGGGCTTCCTGGGCGCCGCCGCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTG CAGGCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCAGCACCACCTGCTGCGCCCCATCGA GGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGC TGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAG CTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGAT CTGGAACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACTACACCAACCTGATCT ACACCTGATCGAGGAGAGCCAGAACCAGCAGGAGAACGAGCAGGAGCTGCTGGAGCT GGACAAGTGGGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGA TCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCG TGAACCGCGTGCGCCAGGGCTACAGCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCCGCG GCCCGACGCCCGAGGGCATCGAGGAGGAGGGCGGCGAGCGCGACCGCAGCAG CCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAG CGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGA ATCGAGGTGGCCCAGCGCATCGCCGCCGCCTTCCTGCACATCCCCCGCCGCATCCGCCAGGGC TTCGAGCGCCCCTGCTGTAACTCGAG

SEQ ID NO:17 GLN422-TYR435B

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG ACCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC AGAAGGAGTACGCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGA GCCCATCCCATCCACTACTGCGCCCCGCCGCTTCGCCATCCTGAAGTGCAACGACAAGAA GTTCAACGCAGCGCCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCC CCGTGGTGAGCACCCAGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC AGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT CAACTGCACCGCCCAACACACACCCGCAAGAGCATCACCATCGGCCCCGGCCGCCCT TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG AAGTGGAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGAC CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG GCCCCAACACCCCAACGGCACCATCACCCTGCCCTGCCGCATCAAGCAGGCCCCCTACGCCC CCCCATCCGCGGCCAGATCCGCTGCAGCAGCACATCACCGGCCTGCTGCTGACCCGCGACG GCGCCAAGGAGATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGAC AACTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCC CACCAAGGCCAAGCGCGCGTGGTGCAGCGCGAGAAGCGCGCGTGACCCTGGGCGCCATGT TCCTGGGCTTCCTGGGCGCCGCCGCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGC AGGCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCAGCAACCTGCTGCGCGCCCATCGAG GCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCT GGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGC TGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATC TGGAACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACTACACCAACCTGATCTA CACCTGATCGAGGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTG GACAAGTGGGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGAT CTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGT GAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCCGCGG CCCGACCGCCCGAGGGCATCGAGGAGGAGGGCGGCGAGCGGACCGCAGCAGCAGC CCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGC TACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCGCATCGTGGAGCTGCTGGGCCGCCGC GGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAA CGAGGTGGCCCAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTT CGAGCGCCCCTGCTGTAACTCGAG

SEQ ID NO:18: LEU122-SER199; ARG426-GLY431

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTTGTGCTGCTGTTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGCCACCACCCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGGGCAACAGCGTGAT CACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGCCCCGCCGG CTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGTGA GCACCGTGCAGTGCACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGCAACGGC AGCCTGGCCGAGGGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCAT CATCGTGCAGCTGAAGGAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCA AGAGCATCACCATCGGCCCCGGCCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCC GCCAGGCCCACTGCAACATCAGCGGCGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACC AAGCTGCAGGCCCAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCC CGAGATCGTGATGCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCT GTTCAACAGCACCTGGAACAACACCATCGGCCCCAACAACACCAACGGCACCATCACCCTGC $\tt CCTGCCGCATCAAGCAGATCATCAACCGCGGCGGCGGCAAGGCCATGTACGCCCCCCCATCC$ GCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAG GAGATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGACAACTGGCG CAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCACCAAGG CCAAGCGCCGCGTGGTGCAGCGCGAGAAGCGCCCGTGACCCTGGGCGCCATGTTCCTGGGC TTCCTGGGCGCCGGCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGAGCGCATCGTGCAGCAGCAGCACCACCTGCTGCGCGCCATCGAGGCCCAGCA GCACCTGCTGCAGCTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGG AGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGC ACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAA CATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACTACACCAACCTGATCTACACCCTGA TCGAGGAGCCAGAACCAGCAGGAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTG GGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCAT GATCGTGGCGCCTGGTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGT GCGCCAGGGCTACAGCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGCGGCCCGACCG ACGCCTGCTGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCC TGCGCGACCTGATCCTGATCGCCGCCCGCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGG CCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTG CAGCGCATCGGCCGCCTTCCTGCACATCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCC CTGCTGTAACTCGAG

SEQ ID NO:19 LEU122-SER199; ARG426-LYS432

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTTGTGCTGCTGTTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGGGCAACAGCGTGAT CACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCCCATCCCCATCACTACTGCGCCCCGCCGG CTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGTGA GCACCGTGCAGTGCACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGCAGCACGGC AGCCTGGCCGAGGAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCAT CATCGTGCAGCTGAAGGAGGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCA AGAGCATCACCATCGGCCCGGCCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCC GCCAGGCCCACTGCAACATCAGCGGCGAGAAGTGGAACACACCCTGAAGCAGATCGTGACC AAGCTGCAGGCCCAGTTCGGCAACAAGACCATCGTGTTCAAGCAGCAGCGGCGGCGACCC CGAGATCGTGATGCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCT GTTCAACAGCACCTGGAACAACACCATCGGCCCCAACAACACCCAACGGCACCATCACCCTGC CCTGCCGCATCAAGCAGATCATCAACCGCGGCGGCAACAAGGCCATGTACGCCCCCCCATCC GCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAG GAGATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGACAACTGGCG CAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAAGG ${\tt CCAAGCGCCGCGTGGTGCAGCGCGAGAAGCGCGCGTGACCCTGGGCGCCATGTTCCTGGGC}$ TTCCTGGGCGCCGCCGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCAGGCCCGC CAGCTGCTGAGCGCATCGTGCAGCAGCAGCACCACCTGCTGCGCGCCATCGAGGCCCAGCA GCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGG AGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGC ACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAA CATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACTACACCAACCTGATCTACACCCTGA TCGAGGAGAGCCAGAACCAGCAGGAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTG GGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCAT GATCGTGGGCGGCTGGTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGT GCGCCAGGGCTACAGCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGCGCCCCGACCG ACGCCTGCTGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCC TGCGCGACCTGATCCTGATCGCCGCCCGCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGG CCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTG CAGCGCATCGGCCGCCTTCCTGCACATCCCCCGCCACCCCAGGGCTTCGAGCGCGCC **CTGCTGTAACTCGAG**

SEQ ID NO: 20: LEU122-SER199; TRP427-GLY431

GAATTCGCCACCATGGATGCAATGAAGAGGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGGGCAACAGCGTGAT CACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGCCCCGCCGG CTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGTGA GCACCGTGCAGTGCACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGCAACGGC AGCCTGGCCGAGGGGGGGGGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCAT CATCGTGCAGCTGAAGGAGGCGTGGAGATCAACTGCACCCGCCCCAACAACAACACCCGCA AGAGCATCACCATCGGCCCGGCCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCC GCCAGGCCCACTGCAACATCAGCGGCGAGAAGTGGAACACACCCTGAAGCAGATCGTGACC AAGCTGCAGGCCCAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCC CGAGATCGTGATGCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCT GTTCAACAGCACCTGGAACAACACCATCGGCCCCAACACCACCACGGCACCATCACCCTGC CCTGCCGCATCAAGCAGATCATCAACCGCTGGGGCGGCAAGGCCATGTACGCCCCCCCATCC GCGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAG GAGATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGACAACTGCGCGACAACTGGCG CAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCACCAAGG CCAAGCGCCGCGTGCACGCGCGAGAAGCGCCCGTGACCCTGGGCGCCATGTTCCTGGGC TTCCTGGGCGCCGCGGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCAGGCCGC CAGCTGCTGAGCGGCATCGTGCAGCAGCAGAACAACCTGCTGCGCGCCCATCGAGGCCCAGCA GCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGG AGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGC ACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAA CATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACTACACCAACCTGATCTACACCCTGA TCGAGGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTG GGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCAT GATCGTGGGCGGCCTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGT GCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCCGGGCCCCGACCG CCCCGAGGGCATCGAGGAGGGGGGGGGGGGGGGCCGCGGCCGCAGCAGCCCCCTGGTGC ACGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCC TGCGCGACCTGATCCTGATCGCCGCCCGCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGG CCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTG CAGCGCATCGGCCGCCTTCCTGCACATCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCC **CTGCTGTAACTCGAG**

60 / 65

SEQ ID NO:21 LYS121-VAL200; ASN425-LYS432

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGGCCCCCGTGATCACCCA GGCCTGCCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGCCCCGCCGGCTTCGC CATCCTGAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGTGAGCACCG TGCAGTGCACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGG CCGAGGAGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTG CAGCTGAAGGAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGAGCAT CACCATCGGCCCGGCCGCCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGC CCACTGCAACATCAGCGGCGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGC AGGCCCAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATC GTGATGCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAAC CATCAAGCAGATCATCAACGCCCCAAGGCCATGTACGCCCCCCCATCCGCGGCCAGATCCG CTGCAGCAGCACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGAGATCAGCAACA CCACCGAGATCTTCCGCCCGGCGGCGGCGACATGCGCGACAACTGGCGCAGCGAGCTGTAC AAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGCGT GGTGCAGCGCGAGAAGCGCGCCGTGACCCTGGGCGCCATGTTCCTGGGCTTCCTGGGCGCCGC CGGCAGCACCATGGGCGCCGCAGCCTGACCTGACCGTGCAGGCCCGCCAGCTGCTGAGCG GCATCGTGCAGCAGCAGAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAG CTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAA GGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGC CCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAACATGACCTGGATG GAGTGGGAGCGCGAGATCGACAACTACACCAACCTGATCTACACCCTGATCGAGGAGAGCCA GAACCAGCAGGAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGG AACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGC CTGGTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTAC AGCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCCGGGCCCCGACCGCCCCGAGGGCATC GAGGAGGAGGCGGCGACCGCGACCGCAGCCCCCTGGTGCACGGCCTGCTGGC CCTGATCTGGGACGACCTGCGCAGCCTGTTCAGCTACCACCGCCTGCGCGACCTGAT CCTGATCGCCGCCCCGCTGGGGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTG GGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACG CCATCGCCATCGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCCAGCGCATCGGC CGCGCCTTCCTGCACATCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTC GAG

SEQ ID NO:22 VAL120-ILE201; ILE 424-ALA433

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGGGCGGCATCACCCAGGCCTG CCCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGCCCCGCCGGCTTCGCCATCCT GAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGTGAGCACCGTGCAGT GCACCCACGGCATCCGCCCGTGGTGAGCACCCAGCTGCTGAACGGCAGCCTGGCCGAG GAGGGCGTGGTGATCCGCAGCGAGACTTCACCGACAACGCCAAGACCATCATCGTGCAGCT GAAGGAGAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGAGCATCACCA TCGGCCCGGCCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACT GCAACATCAGCGGCGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCC CAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGAT GCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCAC CTGGAACAACACCATCGGCCCCAACAACACCCAACGGCACCATCACCCTGCCCTGCCGCATCA AGCAGATCATCGGCGGCGCCATGTACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGC AACATCACCGGCCTGCTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCGAGAT CTTCCGCCCGGCGGCGCGACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGG TGGTGAAGATCGAGCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGCGTGGTGCAGCGC GAGAAGCGCGCGTGACCCTGGGCGCATGTTCCTGGGCTTCCTGGGCGCCGCCGCCGCAGCACC ATGGGCGCCGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGAGCGGCATCGTGCA GCAGCAGAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGT GGGGCATCAAGCAGCTGCAGGCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAG CTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCC CGAGATCGACAACTACACCAGCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGG AGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGAC ATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTG CGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCCGGGCCCCGACCGCCCCGAGGGCATCGAGGAGGAGGG ACGACCTGCGCAGCCTGTTCCAGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCG $\tt CCCGCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTG$ ${\sf CTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATC}$ GCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCCTTCCT GCACATCCCCGCCGCATCCGCCAGGGCTTCGAGCGCCCCTGCTGTAACTCGAG

SEQ ID NO:23: VAL120-ILE201B; ILE424-ALA433

GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGCCCGGCATCACCCAGGCCTGC AAGTGCAACGACAAGAAGTTCAACGGCAGCGCCCCTGCACCAACGTGAGCACCGTGCAGTG CACCCACGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGAACGGCAGCCTGGCCGAGG AGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTG AAGGAGAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGAGCATCACCAT CGGCCCGGCCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTG CAACATCAGCGGCGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCC AGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGACCCCGAGATCGTGATG CACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACC TGGAACACCATCGGCCCCAACACACCAACGGCACCATCACCCTGCCCTGCCGCATCAA GCAGATCATCGGCGGCCCATGTACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA ACATCACCGGCCTGCTGCTGACCCGCGACGCGCGAAGGAGATCAGCAACACCACCGAGATC TTCCGCCCGGCGGCGGCGACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGGT GGTGAAGATCGAGCCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGCGTGGTGCAGCGCG AGAAGCGCGCGTGACCCTGGGCGCCATGTTCCTGGGCTTCCTGGGCGCCGCCGGCAGCACCA TGGGCGCCGCAGCCTGACCCTGACCGTGCAGGCCGCCAGCTGCTGAGCGGCATCGTGCAG CAGCAGAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTG GGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGC TGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCA GAGATCGACAACTACACCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGA GAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGACA TCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGC GCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCT TCCAGACCCGCTTCCCCGCCCCCGCGGCCCCGACCGCCCCGAGGGCATCGAGGAGGAGGGC GGCGAGCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGA CGACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGC CCGCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGC TGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATC GCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCCCTTCCT GCACATCCCCGCCGCATCCGCCAGGGCTTCGAGCGCCCCTGCTGTAACTCGAG

SEQ ID NO:24 VAL120-THR202; ILE424-ALA433

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGGGCGGCGCCACCCAGGCCTG ${\tt CCCCAAGGTGAGCTTCGAGCCCATCCCATCCACTACTGCGCCCCGGCGGCTTCGCCATCCT}$ GAAGTGCAACGACAAGAAGTTCAACGGCAGCGCCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCCCGTGGTGAGCACCCAGCTGCTGAACGGCAGCCTGGCCGAG GAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCT GAAGGAGAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGAGCATCACCA TCGGCCCCGGCCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACT GCAACATCAGCGGCGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCC CAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGAT GCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCAC CTGGAACAACACCATCGGCCCCAACAACACCAACGGCACCATCACCCTGCCCTGCCGCATCA AGCAGATCATCGGCGGCGCCATGTACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGC AACATCACCGGCCTGCTGACCCGCGACGGCGCAAGGAGATCAGCAACACCACCGAGAT CTTCCGCCCGGCGGCGACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGG TGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAAGGCCAAGCGCCGCGTGGTGCAGCGC GAGAAGCGCGCGTGACCCTGGGCGCCATGTTCCTGGGCTTCCTGGGCGCCGCCGCCAGCACC ATGGGCGCCCGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGAGCGGCATCGTGCA GCAGCAGAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGT GGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAG CTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCC CGAGATCGACAACTACACCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGG AGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGAC CGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGC TTCCAGACCCGCTTCCCCGCCCCCCGCGGCCCCGACCGCCCCGAGGGCATCGAGGAGGAGGG CGGCGAGCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGG ACGACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCG CCCGCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTG CTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATC GCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCGCCTTCCT GCACATCCCCGCCGCATCCGCCAGGGCTTCGAGCGCCCCTGCTGTAACTCGAG

SEQ ID NO:25 VAL127-ASN195

GAATTCGCCACCATGGATGCAATGAAGAGGGCTCTGCTGTTGTGCTGCTGTTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCTGTGCGTG GGGGCAGGGAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCC CATCCCCATCCACTACTGCGCCCCGCCGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTT CAACGCAGCGCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCCCG TGGTGAGCACCCAGCTGCTGCAACGCCAGCCTGGCCGAGGAGGGCGTGGTGATCCGCAGC GAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGATCAA CTGCACCGCCCAACAACACCCGCAAGAGCATCACCATCGGCCCGGCCGCGCCTTCTA CGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAGAAGT GGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGACCATC GTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCGGCGG CGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCGGCCC CAACAACACCAACGGCACCATCACCCTGCCCTGCCGCATCAAGCAGATCATCAACCGCTGGC AGGAGGTGGGCAAGGCCATGTACGCCCCCCCATCCGCGCCAGATCCGCTGCAGCAGCAAC ATCACCGGCCTGCTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCGAGATCTT CCGCCCGGCGCGCGACATGCGCGACAACTGCGCGCAGCGAGCTGTACAAGTACAAGTGG TGAAGATCGAGCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGCGTGGTGCAGCGCGAG AAGCGCGCGTGACCCTGGGCGCCATGTTCCTGGGCTTCCTGGGCGCCGCCGGCAGCACCATG GGCGCCCGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCA GCAGAACAACCTGCTGCGCCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGG GCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTG CTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAG AGATCGACAACTACACCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAG AAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGACAT CAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCG CATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTT CCAGACCCGCTTCCCCGCCCCCGCGGCCCCGACGCCCCGAGGGCATCGAGGAGGAGGGCG GCGAGCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGAC GACCTGCGCAGCCTGTTCCGCTGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCC CGCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGAGGCCCTGAAGTACTGGGGCAACCTGCT GCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCG CCGTGGCCGAGGGCACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCCTTCCTGC ACATCCCCGCCGCATCCGCCAGGGCTTCGAGCGCCCCTGCTGTAACTCGAG

65 / 65

SEQ ID NO:26 VAL127-ASN195; ARG426-GLY431

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGTGAAGCTGACCCCCCTGTGCGTG GGGGCAGGGAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCC CATCCCCATCCACTACTGCGCCCCGCCGCCTTCGCCATCCTGAAGTGCAACGACAAGAAGTT CAACGCAGCGCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCCCG TGGTGAGCACCCAGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGCAGC GAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGATCAA CTGCACCCGCCCAACAACAACACCCGCAAGAGCATCACCATCGGCCCCGGCCGCCTTCTA CGCCACCGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAGAAGT GGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGACCATC GTGTTCAAGCAGAGCAGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCGGCGG CGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCGGCCC CAACAACACCAACGGCACCATCACCCTGCCCTGCCGCATCAAGCAGATCATCAACCGCGGCG GCGGCAAGGCCATGTACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACATCACC GGCCTGCTGACCCGCGACGCGGCAAGGAGATCAGCAACACCACCGAGATCTTCCGCCC CGGGGGCGCGACATGCGCGACACTGGCGCAGCGAGCTGTACAAGTACAAGTGGTGAAG ATCGAGCCCCTGGGCGTGGCCCCCACCAAGGCCAAGCGCCGCGTGGTGCAGCGCGAGAAGCG CGCCGTGACCCTGGGCGCATGTTCCTGGGCTTCCTGGGCGCCGCCGGCAGCACCATGGGCGC CCGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCAGA ACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCCATC AAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGG CATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTGGA GACAACTACACCAACCTGATCTACACCCTGATCGAGGAGGAGCCAGAACCAGCAGGAGAAGAA CGAGCAGGAGCTGCTGGACCAAGTGGGCCAGCCTGTGGAACTGGTTCGACATCAGCA AGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCCTGGTGGGCCTGCGCATCG TGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCCAGA CCCGCTTCCCCGCCCCCGCGCCCCGACCGCCCCGAGGGCATCGAGGAGGAGGGCGGCGAG CGCGACCGCGACCGCAGCCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACGACCTG GTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGCAGTA CTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCGCCGTGG CCGAGGGCACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCCCTTCCTGCACATCC CCCGCCGCATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG

SEQUENCE LISTING

<110> Chiron Corporation

<120> MODIFIED HIV ENV POLYPEPTIDES

<130> 1605.100

<140>

<141>

<160> 26

<170> PatentIn Ver. 2.0

<210> 1

<211> 856

<212> PRT

<213> Human immunodeficiency virus

<400> 1

Met Arg Val Lys Glu Lys Tyr Gln His Leu Trp Arg Trp Gly Trp Arg 1 5 10 15

Trp Gly Thr Met Leu Gly Met Leu Met Ile Cys Ser Ala Thr Glu 20 25 30

Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala 35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu 50 60

Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn 65 70 75 80

Pro Gln Glu Val Val Leu Val Asn Val Thr Glu Asn Phe Asn Met Trp \$90\$

Lys Asn Asp Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp 100 105 110

Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Ser 115 120 125

Leu Lys Cys Thr Asp Leu Lys Asn Asp Thr Asn Thr Asn Ser Ser Ser 130 135 140

Gly Arg Met Ile Met Glu Lys Gly Glu Ile Lys Asn Cys Ser Phe Asn 145 \$150\$

Ile Ser Thr Ser Ile Arg Gly Lys Val Gln Lys Glu Tyr Ala Phe Phe 165 170 175

Tyr Lys Leu Asp Ile Ile Pro Ile Asp Asn Asp Thr Thr Ser Tyr Lys 180 185 190

Leu Thr Ser Cys Asn Thr Ser Val Ile Thr Gln Ala Cys Pro Lys Val 195 200205

- Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala 210 $\,$ 215 $\,$ 220 $\,$
- Ile Leu Lys Cys Asn Asn Lys Thr Phe Asn Gly Thr Gly Pro Cys Thr 225 230 235 240
- Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro Val Val Ser 245 250 255
- Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Val Val Ile 260 265 270
- Arg Ser Val Asn Phe Thr Asp Asn Ala Lys Thr Ile Ile Val Gln Leu 275 280 285
- Asn Thr Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg 290 295 300
- Lys Arg Ile Arg Ile Gln Arg Gly Pro Gly Arg Ala Phe Val Thr Ile 305 310 315 320
- Gly Lys Ile Gly Asn Met Arg Gln Ala His Cys Asn Ile Ser Arg Ala 325 330 335
- Lys Trp Asn Asn Thr Leu Lys Gln Ile Ala Ser Lys Leu Arg Glu Gln 340 345 350
- Phe Gly Asn Asn Lys Thr Ile Ile Phe Lys Gln Ser Ser Gly Gly Asp 355 360 365
- Pro Glu Ile Val Thr His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr 370 375 380
- Cys Asn Ser Thr Gln Leu Phe Asn Ser Thr Trp Phe Asn Ser Thr Trp 385 390 395 400
- Ser Thr Glu Gly Ser Asn Asn Thr Glu Gly Ser Asp Thr Ile Thr Leu 405 410 415
- Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Lys Val Gly Lys
 420 425 430
- Ala Met Tyr Ala Pro Pro Ile Ser Gly Gln Ile Arg Cys Ser Ser Asn 435 440445
- Ile Thr Gly Leu Leu Thr Arg Asp Gly Gly Asn Ser Asn Asn Glu 450 455 460
- Ser Glu Ile Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg 465 470 475 480
- Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val
- Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys Arg Ala 500 505 510

2

Val Gly Ile Gly Ala Leu Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Met Thr Leu Thr Val Gln Ala Arg Gln Leu 535 Leu Ser Gly Ile Val Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu 555 550 Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu 570 Gln Ala Arg Ile Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu 585 Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Ala Val Pro Trp Asn Ala Ser Trp Ser Asn Lys Ser Leu Glu Gln Ile Trp Asn His Thr Thr Trp Met Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Ser 630 Leu Ile His Ser Leu Ile Glu Glu Ser Gln Asn Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp 665 Phe Asn Ile Thr Asn Trp Leu Trp Tyr Ile Lys Leu Phe Ile Met Ile Val Gly Gly Leu Val Gly Leu Arg Ile Val Phe Ala Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr His 710 715 Leu Pro Thr Pro Arg Gly Pro Asp Arg Pro Glu Gly Ile Glu Glu 730 Gly Glu Arg Asp Arg Asp Arg Ser Ile Arg Leu Val Asn Gly Ser Leu Ala Leu Ile Trp Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr 760 His Arg Leu Arg Asp Leu Leu Leu Ile Val Thr Arg Ile Val Glu Leu Leu Gly Arg Arg Gly Trp Glu Ala Leu Lys Tyr Trp Trp Asn Leu Leu 790 795 Gln Tyr Trp Ser Gln Glu Leu Lys Asn Ser Ala Val Ser Leu Leu Asn 810 Ala Thr Ala Ile Ala Val Ala Glu Gly Thr Asp Arg Val Ile Glu Val 825

3

Val Gln Gly Ala Cys Arg Ala Ile Arg His Ile Pro Arg Arg Ile Arg 835 840 845

Gln Gly Leu Glu Arg Ile Leu Leu 850 855

<210> 2

<211> 847

<212> PRT

<213> Human immunodeficiency virus

<400> 2

Gly Thr Leu Leu Leu Gly Met Leu Met Ile Cys Ser Ala Val Glu Lys \$20\$

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr 35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val
50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro 65 70 75 80

Gln Glu Ile Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys \$90\$ 95

Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp 100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
115 120 125

His Cys Thr Asn Leu Lys Asn Ala Thr Asn Thr Lys Ser Ser Asn Trp

Lys Glu Met Asp Arg Gly Glu Ile Lys Asn Cys Ser Phe Lys Val Thr

Thr Ser Ile Arg Asn Lys Met Gln Lys Glu Tyr Ala Leu Phe Tyr Lys 165 170 175

Leu Asp Val Val Pro Ile Asp Asn Asp Asn Thr Ser Tyr Lys Leu Ile 180 185 190

Asn Cys Asn Thr Ser Val Ile Thr Gln Ala Cys Pro Lys Val Ser Phe 195 200 205

Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu 210 215 220

Lys Cys Asn Asp Lys Lys Phe Asn Gly Ser Gly Pro Cys Thr Asn Val 225 230 235 240

4

Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Gly Val Val Ile Arg Ser 265 Glu Asn Phe Thr Asp Asn Ala Lys Thr Ile Ile Val Gln Leu Lys Glu 280 Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Thr Ile Gly Pro Gly Arg Ala Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn Ile Ser Gly Glu Lys Trp Asn Asn Thr Leu Lys Gln Ile Val Thr Lys Leu Gln Ala Gln Phe Gly Asn 345 Lys Thr Ile Val Phe Lys Gln Ser Ser Gly Gly Asp Pro Glu Ile Val 360 Met His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn Ser Thr Gln Leu Phe Asn Ser Thr Trp Asn Asn Thr Ile Gly Pro Asn Asn Thr 390 395 Asn Gly Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Arg 405 Trp Gln Glu Val Gly Lys Ala Met Tyr Ala Pro Pro Ile Arg Gly Gln Ile Arg Cys Ser Ser Asn Ile Thr Gly Leu Leu Thr Arg Asp Gly 440 Gly Lys Glu Ile Ser Asn Thr Thr Glu Ile Phe Arg Pro Gly Gly 455 Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Arg Arg Val 490 Val Gln Arg Glu Lys Arg Ala Val Thr Leu Gly Ala Met Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Arg Ser Leu Thr Leu 520 515 Thr Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr 555

Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Val Glu Arg 565 570 575

- Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys 580 585 590
- Leu Ile Cys Thr Thr Ala Val Pro Trp Asn Ala Ser Trp Ser Asn Lys 595 600 605
- Ser Leu Asp Gln Ile Trp Asn Asn Met Thr Trp Met Glu Trp Glu Arg
 610 620
- Glu Ile Asp Asn Tyr Thr Asn Leu Ile Tyr Thr Leu Ile Glu Glu Ser 625 630 635 640
- Gln Asn Gln Glu Lys Asn Glu Glu Leu Leu Glu Leu Asp Lys
 645 650 655
- Trp Ala Ser Leu Trp Asn Trp Phe Asp Ile Ser Lys Trp Leu Trp Tyr
 660 665 670
- Ile Lys Ile Phe Ile Met Ile Val Gly Gly Leu Val Gly Leu Arg Ile 675 680 685
- Val Phe Thr Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr Ser 690 695 700
- Pro Leu Ser Phe Gln Thr Arg Phe Pro Ala Pro Arg Gly Pro Asp Arg 705 710 715 720
- Pro Glu Gly Ile Glu Glu Gly Gly Glu Arg Asp Arg Ser 725 730 735
- Ser Pro Leu Val His Gly Leu Leu Ala Leu Ile Trp Asp Asp Leu Arg
 740 745 750
- Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp Leu Ile Leu Ile
 755 760 765
- Ala Ala Arg Ile Val Glu Leu Leu Gly Arg Arg Gly Trp Glu Ala Leu 770 775 780
- Lys Tyr Trp Gly Asn Leu Leu Gln Tyr Trp Ile Gln Glu Leu Lys Asn 785 790 795 800
- Ser Ala Val Ser Leu Phe Asp Ala Ile Ala Ile Ala Val Ala Glu Gly 805 810 815
- Thr Asp Arg Ile Ile Glu Val Ala Gln Arg Ile Gly Arg Ala Phe Leu 820 825 830
- His Ile Pro Arg Arg Ile Arg Gln Gly Phe Glu Arg Ala Leu Leu 835 840 845

<210> 3

<211> 2310

<212> DNA

<213> Artificial Sequence

```
<220>
<223> Description of Artificial Sequence: Val120-Ala204
<400> 3
gaattegeea eeatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
geagtetteg titegeecag egeegtggag aagetgtggg tgacegtgta etaeggegtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accqaqqtqc acaacqtqtq qqccacccac gcctgcgtqc ccaccqaccc caacccccaq 240
gagategtge tggaqaaegt gaeegagaae tteaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgggcgcc 360
ggegeetgee ceaaggtgag ettegageee ateceeatee actactgege eeeeggeegge 420
ttegecatee tqaaqtqcaa egacaagaag tteaacggca geggeceetg caccaacgtg 480
ageaccgtgc agtgcaccca cggcatccgc cccgtggtga gcacccagct gctgctgaac 540
ggcagcctgg ccgaggaggg cgtggttgatc cgcagcgaga acttcaccga caacgccaag 600
accatcatcg tgcagctgaa ggagagcgtg gagatcaact gcacccgccc caacaacaac 660
accegeaaga geateaceat eggeeeegge egegeettet aegeeacegg egacateate 720
ggcgacatec gecaggeea etgeaacate ageggegaga agtggaacaa caccetgaag 780
cagategtga ceaagetgea ggeceagtte ggeaacaaga ceategtgtt caageagage 840
ageggeggeg acceegagat egtgatgeac agetteaact geggeggega gttettetae 900
tgcaacagca cccagctgtt caacagcacc tggaacaaca ccatcggccc caacaacacc 960
aacggcacca tcaccctgcc ctgccgcatc aagcagatca tcaaccgctg gcaggaggtg 1020
ggcaaggcca tgtacqcccc ccccatccgc ggccagatcc gctgcagcag caacatcacc 1080
ggcctqctqc tqacccqcqa cggcggcaag gagatcagca acaccaccga gatcttccqc 1140
cccggcggcg gcgacatgcg cgacaactgg cgcagcgagc tgtacaagta caaggtggtq 1200
aagategage ceetgggegt ggeeccaace aaggeeaage geegegtggt geagegegag 1260
aagcqcqccq tqaccctqqq cqccatgttc ctgggcttcc tgggcqccqc cqqcaqcacc 1320
atgggcgccc gcagcctgac cctgaccqtg caggcccgcc agctgctgag cggcatcqtg 1380
caqcaqcaqa acaacctqct qcqcqccatc qaqqcccaqc aqcacctqct qcaqctqacc 1440
gtgtggggca tcaagcagct gcaggcccgc gtgctggccg tggagcgcta cctgaaggac 1500
cagcagetge tgggcatetg gggctgcage ggcaagetga tetgcaccae egeegtgcee 1560
tggaacgcca gctggagcaa caagagcctg gaccagatct ggaacaacat gacctggatg 1620
gagtgggage gegagatega caactacace aacetgatet acaccetgat egaggagage 1680
cagaaccagc aggagaagaa cgagcaggag ctgctggagc tggacaagtg ggccagcctg 1740
tggaactggt tcgacatcag caagtggctg tggtacatca agatcttcat catgatcgtg 1800
qqcqqcctqq tqqqcctqcq catcqtqttc accqtqctqa gcatcqtqaa ccqcqtqcqc 1860
cagggetaca geoccetqag ettecagace egetteeeeg eeccegegg eeccgacege 1920
cccgagggca tcqaggagga gggggggag cgcgaccgcg accgcagcag ccccctggtg 1980
cacggcctgc tggccctgat ctgggacgac ctgcgcagcc tgtgcctgtt cagctaccac 2040
cgcctgcgcg acctgatect gategcegec cgcategtgg agetgctggg ccgccgcggc 2100
tgggaggccc tgaagtactg gggcaacctg ctgcagtact ggatccagga gctgaagaac 2160
agegeegtga geetgttega egecategee ategeegtgg eegagggeac egacegeate 2220
ategaqqtqq cccaqeqeat eggeegege tteetgeaca teeceegeeg cateegeeaq 2280
ggcttcgagc gcgccctgct gtaactcgag
<210> 4
<211> 2316
<212> DNA
<213> Artificial Sequence
<223> Description of Artificial Sequence: Vall20-Ile201
<400> 4
gaattegeca ccatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtetteg tttegcccag egcegtggag aagetgtggg tgacegtgta etaeggegtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcqtqc tqqaqaacqt qaccqagaac ttcaacatqt qqaaqaacaa catqqtqqaq 300
cagatqcacq aggacatcat cagcctqtqq qaccaqaqcc tqaaqccctq cqtqqqcqqc 360
```

```
atcacccagg cetgececaa ggtgagette gageceatee ceatecacta etgegeecee 420
geeggetteg ccateetgaa gtgcaacgae aagaagttea aeggeagegg cceetgeace 480
aacgtgagca ccgtgcagtg cacccacggc atccgccccg tggtgagcac ccagctgctg 540
ctgaacggca gcctggccga ggagggcgtg gtgatccgca gcgagaactt caccgacaac 600
gccaagacca tcatcgtgca gctgaaggag agcgtggaga tcaactgcac ccgcccaac 660
aacaacacc gcaagagcat caccatcggc cccggccgcg ccttctacgc caccggcqac 720
atcateggeg acateegeea ggeceactge aacateageg gegagaagtg gaacaacace 780
ctgaagcaga tcgtgaccaa gctgcaggcc cagttcggca acaagaccat cgtgttcaag 840
cagageageg geggegaeee egagategtg atgeaeaget teaactgegg eggegagtte 900
ttctactgca acagcaccca gctgttcaac agcacctgga acaacaccat cggccccaac 960
aacaccaacg gcaccatcac cctgccctgc cgcatcaagc agatcatcaa ccgctggcag 1020
gaggtgggca aggccatgta cgccccccc atccgcggcc agatccgctg cagcagcaac 1080
atcaccggcc tgctgctgac ccgcgacggc ggcaaggaga tcagcaacac caccgagatc 1140
ttccgccccg gcggcggcga catgcgcgac aactggcgca gcgagctgta caagtacaag 1200
gtggtgaaga tegageeect gggegtggee eecaceaagg ceaagegeeg egtggtgeag 1260
cgcgagaagc gcgccgtgac cctgggcgcc atgttcctgg gcttcctggg cgccgccggc 1320
agcaccatgg gcgcccgcag cctgaccctg accgtgcagg cccgccagct gctgagcggc 1380
atcgtgcagc agcagaacaa cctgctgcgc gccatcgagg cccagcagca cctgctgcag 1440
ctgaccgtgt ggggcatcaa gcagctgcag gcccgcgtgc tggccgtgga gcgctacctg 1500
aaggaccagc agetgetggg catetggggc tgcageggca agetgatetg caccaccgcc 1560
gtgccctgga acgccagctg gagcaacaag agcctggacc agatctggaa caacatgacc 1620
tggatggagt gggagcgcga gatcgacaac tacaccaacc tgatctacac cctgatcgag 1680
gagagccaga accagcagga gaagaacgag caggagctgc tggagctgga caagtgggcc 1740
agcctgtgga actggttcga catcagcaag tggctgtggt acatcaagat cttcatcatq 1800
atcgtgggcg gcctggtggg cctgcgcatc gtgttcaccg tgctgagcat cgtgaaccgc 1860
gtgcgccagg gctacagccc cctgagcttc cagacccgct tccccgcccc ccgcggcccc 1920
gaccgccccg agggcatcga ggaggaggcc ggcgagcgcg accgcgaccg cagcagcccc 1980
ctggtgcacg gcctgctggc cctgatctgg gacgacctgc gcagcctgtg cctgttcagc 2040
taccaccqcc tgcgcgacct gatcctgatc gccgcccgca tcgtggagct gctgggccqc 2100
cgeggetggg aggecetgaa gtactgggge aacetgetge agtactggat ecaggagetg 2160
aagaacagcg ccgtgagcct gttcgacgcc atcgccatcg ccgtggccga gggcaccgac 2220
cgcatcatcg aggtggccca gcgcatcggc cgcgccttcc tgcacatccc ccgccgcatc 2280
cgccagggct tcgagcgcgc cctgctgtaa ctcgag
<210> 5
<211> 2322
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Val120-Ile201B
<400> 5
gaattegeca ccatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtcttcq tttcqcccaq cqccqtqqaq aaqctqtqqq tqaccqtqta ctacqqcqtq 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgeacg aggacateat cageetgtgg gaccagagee tgaageeetg egtgeeegge 360
atcacccagg cctgccccaa ggtgagcttc gagcccatcc ccatccacta ctgcgccccc 420
gcoggetteg ceatectgaa gtgcaacgac aagaagttca acggcagegg cecetgcace 480
aacgtgagca ccqtqcagtg cacccacggc atccqccccq tqqtqaqcac ccaqctqctq 540
ctgaacggca gcctggccga ggagggcgtg gtgatccqca gcgagaactt caccgacaac 600
gccaagacca tcatcgtgca gctgaaggag agcgtggaga tcaactgcac ccgccccaac 660
aacaacacc gcaagagcat caccategge eeeggeegeg cettetaege caceggegac 720
atcatcggcg acatccgcca ggcccactgc aacatcagcg gcgagaagtg gaacaacacc 780
ctgaagcaga tcgtgaccaa gctgcaggcc cagttcggca acaagaccat cgtgttcaag 840
cagagcagcg gcggcgaccc cgagatcgtg atgcacagct tcaactgcgg cggcgagttc 900
ttotactgca acagcaccca gctgttcaac agcacctgga acaacaccat cggccccaac 960
```

```
aacaccaacg gcaccatcac cctgccctgc cgcatcaagc agatcatcaa ccgctggcag 1020
gaggtgggca aggccatgta cgccccccc atccgcggcc agatccgctg cagcagcaac 1080
atcaccggcc tgctgctgac ccgcgacggc ggcaaggaga tcagcaacac caccgagatc 1140
ttccqccccq qcggcggcga catgcgcgac aactggcgca gcgagctgta caagtacaag 1200
gtggtgaaga tegageeect gggegtggee cecaceaagg ceaagegeeg egtggtgeag 1260
egegagaage gegeegtgae eetgggegee atgtteetgg getteetggg egeegeegge 1320
ageaccatgg gegeegeag ectgacectg accgtgeagg eccgceaget getgagegge 1380
atogtgcago agoagaacaa ootgotgogo gocatogagg cocagoagca ootgotgcag 1440
ctgaccgtgt ggggcatcaa gcagctgcag gcccgcgtgc tggccgtgga gcgctacctg 1500
aaggaccagc agetgetggg catetgggge tgeageggea agetgatetg caccacegee 1560
gtgccctgga acgccagctg gagcaacaag agcctggacc agatctggaa caacatgacc 1620
tggatggagt gggagegega gatcgacaac tacaccaacc tgatctacac cctgatcgag 1680
gagagccaga accagcagga gaagaacgag caggagctgc tggagctgga caagtgggcc 1740
agcctgtgga actggttcga catcagcaag tggctgtggt acatcaagat cttcatcatg 1800
atcgtgggcg gcctggtggg cctgcgcatc gtgttcaccg tgctgagcat cgtgaaccgc 1860
gtgegecagg getacagece cetgagette cagaceeget teecegecee eegeggecee 1920
gaccgccccg agggcatcga ggaggagggc ggcgagcgcg accgcgaccg cagcagcccc 1980
ctggtgcacg gcctgctggc cctgatctgg gacgacctgc gcagcctgtg cctgttcagc 2040
taccaccgcc tgcgcgacct gatcctgatc gccgcccgca tcgtggagct gctgggccgc 2100
cqcqqctqqq aqqcctgaa gtactggggc aacctgctgc agtactggat ccaggagctg 2160
aagaacagcg ccgtgagcct gttcgacgcc atcgccatcg ccgtggccga gggcaccgac 2220
cgcatcatcg aggtggccca gcgcatcggc cgcgccttcc tgcacatccc ccgccgcatc 2280
cgccagggct tcgagcgcgc cctgctgtaa ctcgagcgtg ct
                                                                  2322
<210> 6
<211> 2328
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Lys121-Val200
<400> 6
gaattogoca coatggatgo aatgaagaga gggototgot gtgtgotgot gotgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccqtqtqqa aqqaqqccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagategtge tggagaaegt gaeegagaae tteaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaaggcc 360
cccgtgatca cccaggcctg ccccaaggtg agettegage ccatecccat ccaetactgc 420
gccccgccg gcttcgccat cctgaagtgc aacgacaaga agttcaacgg cagcggcccc 480
tgcaccaacg tgagcaccgt gcagtgcacc cacggcatcc gccccgtggt gagcacccag 540
ctqctqctqa acqqcaqcct ggccgaggag ggcgtggtga tccgcagcga gaacttcacc 600
gacaacqcca aqaccatcat cgtgcagctg aaggagagcg tggagatcaa ctgcacccgc 660
cccaacaaca acaccegcaa gagcatcacc ateggeceeg geegegeett etaegecace 720
ggcgacatca tcggcgacat ccgccaggcc cactgcaaca tcagcggcga gaagtggaac 780
aacaccctga agcagatcgt gaccaagctg caggcccagt tcggcaacaa gaccatcgtg 840
ttcaagcaga gcagcgggg cgaccccgag atcgtgatgc acagcttcaa ctgcggcggc 900
gagttettet actgeaacag caccagetg tteaacagea eetggaacaa caccategge 960
cccaacaaca ccaacggcac catcaccctg ccctgccgca tcaagcagat catcaaccgc 1020
tggcaggagg tgggcaaggc catgtacgcc cccccatcc geggccagat ccgctgcagc 1080
agcaacatca coqqootgot gotgaccogo gacggoggca aggagatcag caacaccaco 1140
gagatettee geeeeggegg eggegacatg egegacaaet ggegeagega getgtacaag 1200
tacaaggtgg tgaagatega geeectggge gtggeeecca ecaaggeeaa gegeegegtg 1260
gtgcagcgcg agaagcgcgc cgtgaccctg ggcgccatgt tcctgggctt cctgggcgcc 1320
geoggeagea coatgggege eegeageetg accetgaceg tgeaggeeeg ceagetgetg 1380
ageggcateq tgcagcagca gaacaacetg etgegegeca tegaggeeca geageacetg 1440
ctgcagctga ccqtqtqqqq catcaagcag ctgcaggccc gcgtgctggc cgtqqaqcqc 1500
tacctqaaqq accaqcaqct gctgggcatc tggggctgca gcggcaagct gatctgcacc 1560
```

```
accgccgtgc cctggaacgc cagctggagc aacaagagcc tggaccagat ctggaacaac 1620
atgacctgga tggagtggga gcgcgagatc gacaactaca ccaacctgat ctacaccctg 1680
atcgaggaga gccagaacca gcaggagaag aacgagcagg agctgctgga gctggacaag 1740
tgggccagcc tgtggaactg gttcgacatc agcaagtggc tgtggtacat caagatcttc 1800
atcatgateg tgggeggeet ggtgggeetg egeategtgt teaeegtget gageategtg 1860
aaccgcgtgc gccagggcta cagccccctg agcttccaga cccgcttccc cgccccccgc 1920
ggccccgacc gccccgaggg catcgaggag gagggcggcg agcgcgaccg cgaccgcagc 1980
agecccetgg tgcacggcct gctggccctg atctgggacg acctgcgcag cctgtgcctg 2040
ttcagctacc accgcctgcg cgacctgatc ctgatcgccg cccgcatcgt ggagctgctg 2100
ggccgccgcg gctgggaggc cctgaagtac tggggcaacc tgctgcagta ctggatccag 2160
gagetgaaga acagegeegt gageetgtte gaegecateg ceategeegt ggeegaggge 2220
accgaccgca teategaggt ggeccagege ateggeegeg cetteetgca catececege 2280
cgcatccgcc agggcttcga gcgcgccctg ctgtaactcg agcgtgct
<210> 7
<211> 2334
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Leu122-Ser199
<400> 7
gaattegeea eeatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
geagtetteg tttegeceag egeegtggag aagetgtggg tgaeegtgta etaeggegtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagategtge tggagaaegt gaeegagaae tteaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcetgtgg gaccagagee tgaageeetg egtgaagetg 360
ggcaacageg tgatcaceca ggcctgcccc aaggtgaget tegageecat eeccatecae 420
tactgcgccc ccgccggctt cgccatcctg aagtgcaacg acaagaagtt caacggcagc 480
ggoccotgoa coaacgtgag caccgtgcag tgcacccacg gcatccgccc cgtggtgagc 540
acceagetge tgetgaaegg cageetggee gaggaggeg tggtgateeg cagegagaae 600
ttcaccgaca acgccaagac catcatcgtg cagctgaagg agagcgtgga gatcaactgc 660
accegecea acaacaacae eegcaagage atcaccateg geeceggeeg egeettetae 720
gccaccggcg acatcatcgg cgacatccgc caggcccact gcaacatcag cggcgagaag 780
tggaacaaca ccctgaagca gatcgtgacc aagctgcagg cccagttcgg caacaagacc 840
atogtgttca agcagagcag cggcggcgac cccgagatcg tgatgcacag cttcaactgc 900
ggcggcgagt tettetactg caacagcace cagetgttea acageacetg gaacaacace 960
ateggeecca acaacaccaa eggeaccate accetgeect geegeateaa geagateate 1020
aaccgctggc aggaggtggg caaggccatg tacgccccc ccatccgcgg ccagatccgc 1080
tgcagcagca acatcacegg cetgetgetg accegegaeg geggcaagga gatcagcaac 1140
accaccgaga tetteegeee eggeggegge gacatgegeg acaactggeg cagegagetg 1200
tacaaqtaca aqqtqqtqaa gatcqagccc ctgggcqtgg cccccaccaa ggccaagcgc 1260
egegtggtge agegegagaa gegegeegtg accetgggeg ceatgtteet gggetteetg 1320
ggegeegeeg geageaceat gggegeege ageetgaeee tgaeegtgea ggeeegeeag 1380
ctgctgagcg gcatcgtgca gcagcagaac aacctgctgc gcgccatcga ggcccagcag 1440
cacctgctgc agctgaccgt gtggggcatc aagcagctgc aggcccgcgt gctggccgtg 1500
gagegetace tgaaggacea geagetgetg ggeatetggg getgeagegg caagetgate 1560
tgcaccaccg ccgtgccctg gaacgccagc tggagcaaca agagcctgga ccagatctgg 1620
aacaacatga cctggatgga gtgggagcgc gagatcgaca actacaccaa cctgatctac 1680
accetgateg aggagageea gaaccageag gagaagaaeg ageaggaget getggagetg 1740
gacaagtggg ccagcetgtg gaactggttc gacatcagca agtggctgtg gtacatcaag 1800
atetteatea tgategtggg eggeetggtg ggeetgegea tegtgtteae egtgetgage 1860
atogtgaacc gogtgogcca gggctacagc cocctgagct tocagacccg cttcccogcc 1920
ccccgcggcc ccgaccgccc cgagggcatc gaggaggagg gcggcgagcg cgaccgcgac 1980
cgcagcagcc ccctggtgca cggcctgctg gccctgatct gggacgacct gcgcagcctg 2040
tgcctgttca qctaccaccg cctgcgcgac ctgatcctga tcgccgcccg catcgtggag 2100
ctgctgggcc gccgcggctg ggaggccctg aagtactggg gcaacctgct gcagtactgg 2160
```

```
atccaggage tgaagaacag egeegtgage etgttegaeg ceategeeat egeegtggee 2220
qaqqqcaccq accqcatcat cgaggtggcc cagcgcatcg gccgcgcctt cctgcacatc 2280
cecegeegea teegeeaggg ettegagege geeetgetgt aactegageg tget
<210> 8
<211> 2316
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Val120-Thr202
<400> 8
gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcaqtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caaccccag 240
gagatogtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcetgtgg gaccagagee tgaageeetg egtgggegge 360
qccacccaqq cctqccccaa ggtgagcttc gagcccatcc ccatccacta ctgcgccccc 420
geoggetteg ccatectgaa gtgcaacgac aagaagttca acggcagegg cccctgcacc 480
aacqtqaqca ccqtqcaqtq cacccacggc atccgccccq tggtgagcac ccagctgctg 540
ctqaacqqca qcctgqccga ggagggcgtg gtgatccgca gcgagaactt caccgacaac 600
gccaagacca tcatcgtgca gctgaaggag agogtggaga tcaactgcac ccgccccaac 660
aacaacacc gcaagagcat caccateggc cceggcegeg cettetacgc caceggegac 720
atcateggeg acateegeea ggeecactge aacateageg gegagaagtg gaacaacace 780
ctgaagcaga tcgtgaccaa gctgcaggcc cagttcggca acaagaccat cgtgttcaag 840
caqaqcaqcq qcqqcqaccc cgagatcgtg atgcacagct tcaactgcgg cggcgagttc 900
ttctactqca acagcaccca gctgttcaac agcacctgga acaacaccat cggccccaac 960
aacaccaacg gcaccatcac cctgccctgc cgcatcaagc agatcatcaa ccgctggcag 1020
gaggtgggca aggccatgta cgccccccc atccgcggcc agatccgctg cagcagcaac 1080
atcaccqqcc tqctqctgac ccgcgacggc ggcaaggaga tcagcaacac caccgagatc 1140
ttccgccccg gcggcggcga catgcgcgac aactggcgca gcgagctgta caagtacaag 1200
gtggtgaaga tcgagcccct gggcgtggcc cccaccaagg ccaagcgccg cgtggtgcag 1260
cgcgagaagc gcgccgtgac cctgggcgcc atgttcctgg gcttcctggg cgccgccggc 1320
agcaccatgg gegecegeag cetgaccetg acegtgeagg ceegecaget getgagegge 1380
atcqtqcaqc aqcaqaacaa cctgctgcgc gccatcgagg cccagcagca cctgctgcag 1440
ctgaccgtgt ggggcatcaa gcagctgcag gcccgcgtgc tggccgtgga gcgctacctg 1500
aaggaccagc agetgetggg catetgggge tgeageggea agetgatetg caccacegee 1560
gtgccctgga acgccagctg gagcaacaag agcctggacc agatctggaa caacatgacc 1620
tggatggagt gggagcgcga gatcgacaac tacaccaacc tgatctacac cctgatcgag 1680
gagagccaga accagcagga gaagaacgag caggagctgc tggagctgga caagtgggcc 1740
agcctgtgga actggttcga catcagcaag tggctgtggt acatcaagat cttcatcatg 1800
atcqtqqqcq qcctqqtqqq cctqcgcatc gtgttcaccq tgctgagcat cgtgaaccgc 1860
qtqcqccaqq qctacaqccc cctgagettc cagacccget tccccgcccc ccgcggcccc 1920
gaccqccccq aqqqcatcga ggaggaggc ggcgagcgcg accgcgaccg cagcagcccc 1980
ctggtgcacg gcctgctggc cctgatctgg gacgacctgc gcagcctgtg cctgttcagc 2040
taccaccgcc tgcgcgacct gatcctgatc gccgcccgca tcgtggagct gctgggccgc 2100
cgcggctggg aggccctgaa gtactggggc aacctgctgc agtactggat ccaggagctg 2160
aagaacagcg ccgtgagcct gttcgacgcc atcgccatcg ccgtggccga gggcaccgac 2220
cgcatcatcg aggtggccca gcgcatcggc cgcgccttcc tgcacatccc ccgccgcatc 2280
cgccagggct tcgagcgcc cctgctgtaa ctcgag
                                                                  2316
<210> 9
<211> 2541
<212> DNA
<213> Artificial Sequence
<220>
```

<223> Description of Artificial Sequence: Trp427-Gly431 gaattegeea ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60 quagtetteq tttegeccag egeegtggag aagetgtggg tgacegtgta etaeggegtg 120 cccqtqtqqa aggagqccac caccacctq ttctqcqcca gcgacqccaa ggcctacqac 180 accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240 gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360 accecetgt gegtgaccet geactgeace aacctgaaga acgccaccaa caccaagage 420 agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480 agcatcegea acaagatgea gaaggagtac gecetgttet acaagetgga egtggtgeee 540 ategacaacg acaacaccag ctacaagetg atcaactgca acaccagegt gatcacccag 600 geotgeecca aggtgagett egageecate eccatecact actgegeece egeeggette 660 gccatcctga agtgcaacga caagaagttc aacggcagcg gcccctgcac caacgtgagc 720 accgtgcagt gcacccacgg catccgcccc gtggtgagca cccagctgct gctgaacggc 780 agectggeeg aggagggegt ggtgateege agegagaact teacegaeaa eqecaaqaee 840 atcatcqtqc aqctqaaqqa qaqcqtggag atcaactgca cccgccccaa caacaacacc 900 cqcaaqaqca tcaccategg ccccggccgc gccttctacg ccaccggcga catcateggc 960 gacatccqcc aggcccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020 atcqtqacca agctqcaggc ccagttcggc aacaagacca tcgtgttcaa gcagagcagc 1080 qqcqqcqacc ccqaqatcqt qatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140 aacaqcaccc aqctqttcaa cagcacctgg aacaacacca tcggccccaa caacaccaac 1200 qqcaccatca ccctqccctq ccqcatcaag cagatcatca accgctgggg cggcaaggcc 1260 atqtacqccc cccccatccg cggccagatc cgctgcagca gcaacatcac cggcctgctg 1320 ctgacccgcg acggcggcaa ggagatcagc aacaccaccg agatcttccg ccccggcggc 1380 qqcqacatqc qcqacaactg gcgcagcgag ctgtacaagt acaaggtggt gaagatcgag 1440 cccctgggcg tggccccac caaggccaag cgccgcgtgg tgcagcgcga gaagcgcgcc 1500 gtgaccetgg gegecatgtt cetgggette etgggegeeg ceggeageac catgggegee 1560 cgcagcctga ccctgaccgt gcaggcccgc cagctgctga gcggcatcgt gcagcagcag 1620 aacaacctgc tgcgcgccat cgaggcccag cagcacctgc tgcagctgac cgtgtggggc 1680 atcaagcagc tgcaggcccg cgtgctggcc gtggagcgct acctgaagga ccagcagctg 1740 ctgggcatct ggggctgcag cggcaagctg atctgcacca ccgccgtgcc ctggaacgcc 1800 agetggagea acaagageet ggaccagate tggaacaaca tgacetggat ggagtgggag 1860 egegagateg acaactacae caacetgate tacaceetga tegaggagag eeagaaceag 1920 caqqaqaaqa acqaqcaqqa gctgctggag ctggacaagt gggccagcct gtggaactgg 1980 ttcgacatca gcaagtggct gtggtacatc aagatcttca tcatgatcgt gggcggcctg 2040 gtgggcctgc gcatcgtgtt caccgtgctg agcatcgtga accgcgtgcg ccagggctac 2100 agccccctga gcttccagac ccgcttcccc gcccccgcg gccccgaccg ccccgagggc 2160 atcgaggagg agggcggcga gcgcgaccgc gaccgcagca gccccctggt gcacggcctg 2220 ctggccctga tctgggacga cctgcgcagc ctgtgcctgt tcagctacca ccgcctgcgc 2280 gacctqatcc tqatcqccgc ccgcatcgtg gagctgctgg gccgccgcgg ctgggaggcc 2340 ctqaaqtact qqqqcaacct gctgcagtac tggatccagg agctgaagaa cagcgccgtg 2400 agectqttcq acqccatcgc catcgccgtg gccgagggca ccgaccgcat catcgaggtg 2460 qeecaqeqea tegqeeqeqe ettectgeae atecceegee geateegeea gggettegag 2520 2541 cgcgccctgc tgtaactcga g <210> 10 <211> 2541 <212> DNA <213> Artificial Sequence <220> <223> Description of Artificial Sequence: Arg426-Gly431 <400> 10 qaattcqcca ccatqqatqc aatqaagaga gggctctgct gtgtgctgct gctgtgtgga 60

gcagtetteg tttegeccag egcegtggag aagetgtggg tgacegtgta etaeggegtg 120 eccqtgtgga aggaggecae caccaccetg ttetgegeca gegacgecaa ggeetaegae 180

```
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
acceccetgt gegtgaccet geactgeace aacetgaaga aegecaceaa caccaagage 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcatccgca acaagatgca gaaggagtac gccctgttct acaagctgga cgtggtgccc 540
atcgacaacg acaacaccag ctacaagctg atcaactgca acaccagcgt gatcacccag 600
geetgeecea aggtgagett egageecate eccatecaet aetgegeece egeeggette 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gcccctgcac caacgtgagc 720
accgtgcagt gcacccacgg catccgcccc gtggtgagca cccagctgct gctgaacggc 780
agcotggeog aggaggogt ggtgatoogo agogagaact toacogacaa cgccaagacc 840
atcategtge agetgaagga gagegtggag atcaactgea ceegeceeaa caacaacace 900
cgcaagagca tcaccatcgg ccccggccgc gccttctacg ccaccggcga catcatcggc 960
gacatccgcc aggcccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
atogtgacca agotgcaggo ccagttoggo aacaagacca togtgttcaa gcagagcago 1080
ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcaccc agctgttcaa cagcacctgg aacaacacca tcggccccaa caacaccaac 1200
ggcaccatca coetgeeetg eegcatcaag cagatcatca acegeggegg eggcaaggee 1260
atgtacgccc cccccatccg cggccagatc cgctgcagca gcaacatcac cggcctgctg 1320
ctgaccogcg acggcggcaa ggagatcagc aacaccaccg agatcttccg ccccggcggc 1380
ggcgacatgc gcgacaactg gcgcagcgag ctgtacaagt acaaggtggt gaagatcgag 1440
cccctgggcg tggcccccac caaggccaag cgccgcgtgg tgcagcgcga gaagcgcgcc 1500
gtgaccetgg gegecatgtt cetgggette etgggegeeg eeggeageac catgggegee 1560
cgcagcctga ccctgaccgt gcaggcccgc cagctgctga gcggcatcgt gcagcagcag 1620
aacaacctgc tgcgcgccat cgaggcccag cagcacctgc tgcagctgac cgtgtggggc 1680
atcaagcagc tgcaggcccg cgtgctggcc gtggagcgct acctgaagga ccagcagctg 1740
ctgggcatct ggggctgcag cggcaagctg atctgcacca ccgccgtgcc ctggaacgcc 1800
agctggagca acaagagcct ggaccagatc tggaacaaca tgacctggat ggagtgggag 1860
cgcgagatcg acaactacac caacctgatc tacaccctga tcgaggagag ccagaaccag 1920
caqqaqaaqa acqaqcagga getgetggag etggacaagt gggecageet gtggaactgg 1980
ttcgacatca gcaagtggct gtggtacatc aagatettca teatgategt gggeggeetg 2040
gtgggcctgc gcatcgtgtt caccgtgctg agcatcgtga accgcgtgcg ccagggctac 2100
agccccctga gcttccagac cogcttcccc gcccccgcg gccccgaccg ccccgagggc 2160
atcgaggagg agggcggcga gcgcgaccgc gaccgcagca gccccctggt gcacggcctg 2220
ctggccctga tctgggacga cctgcgcagc ctgtgcctgt tcagctacca ccgcctgcgc 2280
gacctgatec tgategeege cegeategtg gagetgetgg geegeegegg etgggaggee 2340
ctgaagtact ggggcaacct gctgcagtac tggatccagg agctgaagaa cagcgccgtg 2400
agectgtteg acgecatege categoogtg geogagggca cegacegeat categaggtg 2460
geocagegea teggeegege etteetgeae ateceeegee geateegeea gggettegag 2520
                                                                  2541
cgcgccctgc tgtaactcga g
<210> 11
<211> 2541
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Arg426-Gly431B
<400> 11
gaattegeca ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtqtqqa aqqaqccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatogtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcetgtgg gaccagagee tgaageeetg egtgaagetg 360
accecetty gegtgaccet geactgeace aacetgaaga acgecaceaa caccaagage 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcatccgca acaagatgca gaaggagtac gccctgttct acaagctgga cgtggtgccc 540
```

```
atogacaacg acaacaccag ctacaagctg atcaactgca acaccagcgt gatcacccag 600
gcctgcccca aggtgagctt cgagcccatc cccatccact actgcgcccc cgccggcttc 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gcccctgcac caacgtgagc 720
acceptgeagt geacceaegg catcegeece gtggtgagea eccagetget getgaaegge 780
agcctggccg aggagggcgt ggtgatccgc agcgagaact tcaccgacaa cgccaagacc 840
atcatcqtqc agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacacc 900
cgcaagagca tcaccatcgg ccccggccgc gccttctacg ccaccggcga catcatcggc 960
gacatccgcc aggcccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
ategtgacca agetgeagge ceagttegge aacaagacca tegtgtteaa geagageage 1080
ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacaqcaccc agctgttcaa cagcacctgg aacaacacca tcggccccaa caacaccaac 1200
ggcaccatca ccctgccctg ccgcatcaag cagatcatca accgcggcag cggcaaggcc 1260
atgtacgecc eccecatecg eggecagate egetgeagea geaacateae eggectgetg 1320
ctgacccgcg acggcggcaa ggagatcagc aacaccaccg agatcttccg ccccggcggc 1380
ggcgacatgc gcgacaactg gcgcagcgag ctgtacaagt acaaggtggt gaagatcgag 1440
cccctgggcg tggcccccac caaggccaag cgccgctgg tgcagcgcga gaagcgcgcc 1500
gtgaccetgg gegecatgtt cetgggette etgggegeeg eeggeageac eatgggegee 1560
cgcagcctga ccctgaccgt gcaggcccgc cagctgctga gcggcatcgt gcagcagcag 1620
aacaacetge tgegegeeat egaggeeeag cageacetge tgeagetgae egtgtgggge 1680
atcaagcagc tgcaggcccg cgtgctggcc gtggagcgct acctgaagga ccagcagctg 1740
ctgggcatct ggggctgcag cggcaagctg atctgcacca ccgccgtgcc ctggaacgcc 1800
agetggagea acaagageet ggaccagate tggaacaaca tgacctggat ggagtgggag 1860
cgcgagatcg acaactacac caacctgatc tacaccctga tcgaggagag ccagaaccag 1920
caggagaaga acgagcagga gctgctggag ctggacaagt gggccagcct gtggaactgg 1980
ttcqacatca gcaagtggct gtggtacatc aagatcttca tcatgatcgt gggcggcctg 2040
gtgggcctgc gcatcgtgtt caccgtgctg agcatcgtga accgcgtgcg ccagggctac 2100
ageccetga gettecagae cegetteece gecceegeg geccegaeeg cecegaggge 2160
atcgaggagg agggcggcga gcgcgaccgc gaccgcagca gccccctggt gcacggcctg 2220
ctgqccctga tctgggacga cctgcgcagc ctgtgcctgt tcagctacca ccgcctgegc 2280
gacctgatcc tgatcgccgc ccgcatcgtg gagctgctgg gccgccgcgg ctgggaggcc 2340
ctqaaqtact qqqqcaacct gctgcagtac tggatccagg agctgaagaa cagcgccgtg 2400
agectatteg acaccatege categorateg googaaggea cogacogcat categaaggtg 2460
gcccagegea teggeegege etteetgeae ateceeegee gcateegeea gggettegag 2520
cgcgccctgc tgtaactcga g
<210> 12
<211> 2541
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Arg426-Lys432
<400> 12
gaattegeca ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
quartetteq ttteqeccaq eqeegtggag aagetgtggg tgacegtgta etaeggegtg 120
cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
accccctgt gcgtgaccct gcactgcacc aacctgaaga acgccaccaa caccaagagc 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcatccgca acaagatgca gaaggagtac gccctgttct acaagctgga cgtggtgccc 540
atcgacaacg acaacacag ctacaagctg atcaactgca acaccagcgt gatcacccag 600
gcctgcccca aggtgagctt cgagcccatc cccatccact actgcgcccc cgccggcttc 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gcccctgcac caacgtgagc 720
accgtgcagt gcacccacgg catccgcccc gtggtgagca cccagctgct gctgaacggc 780
agcetggeeg aggagggegt ggtgateege agegagaaet teaecgaeaa egeeaagaee 840
atcatcgtgc agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacacc 900
```

```
cgcaagagca tcaccatcgg ccccggccgc gccttctacg ccaccggcga catcatcggc 960
gacatcegcc aggcccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
ategtgacca agetgeagge ceagttegge aacaagacca tegtgtteaa geagageage 1080
ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcaccc agctgttcaa cagcacctgg aacaacacca tcggccccaa caacaccaac 1200
ggcaccatca ccctgccctg ccgcatcaag cagatcatca accgcggcgg caacaaggcc 1260
atgtacgecc eccecatecg eggecagate egetgeagea geaacateae eggeetgetg 1320
ctgacccgcg acggcggcaa ggagatcagc aacaccaccg agatcttccg ccccggcggc 1380
ggcgacatgc gcgacaactg gcgcagcgag ctgtacaagt acaaggtggt gaagatcgag 1440
cccctgggcg tggccccac caaggccaag cgccgcgtgg tgcagcgcga gaagcgcgcc 1500
gtgaccctgg gcgccatgtt cctgggcttc ctgggcgccg ccggcagcac catgggcgcc 1560
cgcagcctga ccctgaccgt gcaggcccgc cagctgctga gcggcatcgt gcagcagcag 1620
aacaacctgc tgcgcgccat cgaggcccag cagcacctgc tgcagctgac cgtgtggggc 1680
atcaagcagc tgcaggcccg cgtgctggcc gtggagcgct acctgaagga ccagcagctg 1740
ctgggcatct ggggctgcag cggcaagctg atctgcacca ccgccgtgcc ctggaacgcc 1800
agetggagea acaagageet ggaccagate tggaacaaca tgacetggat ggagtgggag 1860
cgcgagatcg acaactacac caacctgatc tacaccctga tcgaggagag ccagaaccag 1920
caggagaaga acgagcagga gctgctggag ctggacaagt gggccagcct gtggaactgg 1980
ttcgacatca gcaagtggct gtggtacatc aagatcttca tcatgatcgt gggcggcctg 2040
gtgggcctgc gcatcgtgtt caccgtgctg agcatcgtga accgcgtgcg ccagggctac 2100
ageccectga gettecagae eegetteeee geeceeegeg geecegaeeg eeeegaggge 2160
ategaggagg agggeggega gegegaeege gaeegeagea geeeeetggt geaeggeetg 2220
ctggccctga tctgggacga cctgcgcagc ctgtgcctgt tcagctacca ccgcctgcgc 2280
gacctgatec tgategeege degeategtg gagetgetgg geogeogegg etgggaggee 2340
ctgaagtact ggggcaacct gctgcagtac tggatccagg agctgaagaa cagcgccgtg 2400
agcetgtteg acgecatege categeegtg geegagggea cegacegeat categaggtg 2460
geccagegea teggeegege etteetgeac atecceegee geateegeea gggettegag 2520
cgcgccctgc tgtaactcga g
<210> 13
<211> 2535
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Asn425-Lys432
<400> 13
gaattegeea ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caaccccag 240
gagategtge tqgagaaegt gaeegagaae tteaacatgt ggaagaacaa catggtggag 300
caqatqcacq aqqacatcat caqcctgtgg gaccagagec tgaagecetg cgtgaagetg 360
accecetqt qeqtqaceet geactgeace aacetgaaga aegecaceaa caccaagage 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcatccgca acaagatgca gaaggagtac gccctgttct acaagctgga cgtggtgccc 540
ategacaacg acaacaccag ctacaagetg ateaactgca acaccagegt gateacccag 600
gcctgcccca aggtgagctt cgagcccatc cccatccact actgcgcccc cgccggcttc 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gcccctgcac caacgtgagc 720
acceptgeagt geacceaegg cateegeece gtggtgagea ceeagetget getgaaegge 780
agectggeeg aggagggegt ggtgateege agegagaaet teaecgaeaa egecaagaee 840
atcateqtqc aqctqaaqqa qaqcqtggag atcaactgca cccqcccaa caacaacacc 900
cgcaagagca tcaccatcgg ccccggccgc gccttctacg ccaccggcga catcatcggc 960
gacateegee aggeeeactg caacateage ggegagaagt ggaacaacae cetgaageag 1020
atogtgacca agotgcaggo coagttoggo aacaagacca togtgttoaa goagagoago 1080
ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcaccc agctgttcaa cagcacctgg aacaacacca teggeeecaa caacaccaac 1200
ggcaccatca ccctqccctg ccgcatcaag cagatcatca acgcccccaa ggccatgtac 1260
```

```
gecececca teegeggeca gateegetge ageageaaca teaceggeet getgetgaee 1320
cgcgacggcg gcaaggagat cagcaacacc accgagatet tecgececgg cggcggcgac 1380
atgogogaca actggcgcag cgagctgtac aagtacaagg tggtgaagat cgagcccttg 1440
ggcgtggccc ccaccaaggc caagcgccgc gtggtgcagc gcgagaagcg cgccgtgacc 1500
ctgggcgcca tgttcctggg cttcctgggc gccgccggca gcaccatggg cgcccgcagc 1560
ctgaccetga cegtgeagge cegecagetg etgageggea tegtgeagea geagaacaac 1620
etgetgegeg ceategagge ceageageae etgetgeage tgaeegtgtg gggeateaag 1680
cagetgeagg ceegegtget ggeegtggag egetacetga aggaceagea getgetggge 1740
atctggggct gcagcggcaa gctgatctgc accaccgccg tgccctggaa cgccagctgg 1800
agcaacaaga geetggacca gatetggaac aacatgacet ggatggagtg ggagcgegag 1860
ategacaact acaccaacet gatetacace etgategagg agagecagaa ecageaggag 1920
aagaacgagc aggagctgct ggagctggac aagtgggcca gcctgtggaa ctggttcgac 1980
atcagcaagt ggctgtggta catcaagatc ttcatcatga tcgtgggcgg cctggtgggc 2040
ctgcgcatcg tgttcaccgt gctgagcatc gtgaaccgcg tgcgccaggg ctacagcccc 2100
ctgagettee agaceegett eccegeecce egeggeeceg acegeecega gggeategag 2160
gaggagggcg gcgagcgca ccgcgaccgc agcagccccc tggtgcacgg cctgctggcc 2220
ctgatctggg acgacctgcg cagcctgtgc ctgttcagct accaccgcct gcgcgacctg 2280
atcctgatcg ccgcccgcat cgtggagctg ctggggccgcc gcggctggga ggccctgaag 2340
tactggggca acctgctgca gtactggatc caggagctga agaacagcgc cgtgagcctg 2400
ttcgacgcca tcgccatcgc cgtggccgag ggcaccgacc gcatcatcga ggtggcccag 2460
cgcatcggcc gegecttect gcacatcccc egeogeatec gccagggett egagegege 2520
ctgctgtaac tcgag
                                                                  2535
<210> 14
<211> 2529
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Ile424-Ala433
<400> 14
gaattegeea ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cageetgtgg gaccagagee tgaageectg egtgaagetg 360
accecctqt qeqtgaccet geactgeace aacctgaaga acgecaceaa caccaagage 420
agcaactgga aggagatgga cegeggegag atcaagaact geagetteaa ggtgaceace 480
agcatcogca acaagatgca gaaggagtac gccctgttct acaagctgga cgtggtgccc 540
ategacaacg acaacaccag ctacaagetg atcaactgca acaccagegt gatcacccag 600
gcctgcccca aggtgagctt cgagcccatc cccatccact actgcgcccc cgccggcttc 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gcccctgcac caacgtgagc 720
accgtgcagt gcacccacgg catccgccc gtggtgagca cccagctgct gctgaacggc 780
agectggeeg aggagggegt ggtgateege agegagaact teacegacaa egecaagace 840
atcategtge agetgaagga gagegtggag atcaaetgea eeegeeecaa caacaacace 900
cgcaagagca tcaccategg ccceggccgc gccttctacg ccaccggcga catcateggc 960
gacateegee aggeeeactg caacateage ggegagaagt ggaacaacae cetgaageag 1020
atcgtgacca agctgcaggc ccagttcggc aacaagacca tcgtgttcaa gcagagcagc 1080
ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcaccc agctgttcaa cagcacctgg aacaacacca tcggccccaa caacaccaac 1200
ggcaccatca ccctqccctg ccgcatcaag cagatcatcg gcggcgccat gtacgccccc 1260
cccatccgcg gccagatccg ctgcagcagc aacatcaccg gcctgctgct gacccgcgac 1320
ggcggcaagg agatcagcaa caccaccgag atcttccgcc ccggcggcgg cgacatgcgc 1380
gacaactggc gcagcgagct gtacaagtac aaggtggtga agatcgagcc cctgggcgtg 1440
gcccccacca aggccaagcg ccgcgtggtg cagcgcgaga agcgcgccgt gaccctgggc 1500
gccatgttcc tgggcttcct gggcgccgcc ggcagcacca tgggcgcccg cagcctgacc 1560
ctgaccgtgc aggcccgcca gctgctgagc ggcatcgtgc agcagcagaa caacctgctg 1620
```

```
egegecateg aggeceagea geacetgetg cagetgaceg tgtggggeat caageagetg 1680
caggeeegeg tgetggeegt ggagegetac etgaaggace ageagetget gggeatetgg 1740
ggctgcagcg gcaagctgat ctgcaccacc gccgtgccct ggaacgccag ctggagcaac 1800
aagagoctgg accagatctg gaacaacatg acctggatgg agtgggagcg cgagatcgac 1860
aactacacca acctgatcta caccetgate gaggagagee agaaccagea ggagaagaac 1920
gagcaggagc tgctggagct ggacaagtgg gccagcctgt ggaactggtt cgacatcagc 1980
aaqtqqctqt ggtacatcaa gatcttcatc atgatcgtgg gcggcctggt gggcctgcgc 2040
atogtgttca ccgtgctgag catcgtgaac cgcgtgcgcc agggctacag ccccctgagc 2100
ttccaqaccc gcttccccgc ccccgcggc cccgaccgcc ccgaggggcat cgaggaggag 2160
ggeggegage gegacegega eegeageage ceeetggtge aeggeetget ggeeetgate 2220
tqqqacqacc tqcqcaqcct gtgcctgttc agctaccacc gcctgcgcga cctgatcctg 2280
ategeegeee geategtgga getgetggge egeegegget gggaggeeet gaagtaetgg 2340
ggcaacctgc tgcagtactg gatccaggag ctgaagaaca gcgccgtgag cctgttcgac 2400
qccatcqcca tcqccqtggc cgagggcacc gaccgcatca tcgaggtggc ccagcgcatc 2460
ggccgcgcct tcctgcacat cccccgccgc atccgccagg gcttcgagcg cgccctgctg 2520
taactcgag
<210> 15
<211> 2523
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Ile423-Met434
<400> 15
gaattegeea ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtetteg tttegeccag egeegtggag aagetgtggg tgacegtgta etaeggegtg 120
cccqtqtqqa aqqaqqccac caccaccctg ttctgcgcca gcgacgccaa ggcctacgac 180
accqaqqtqc acaacqtqtq qqccacccac gcctgcgtgc ccaccgaccc caacccccag 240
qaqatcqtqc tqqaqaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatqcacq aqqacatcat cagcetgtgg gaccagagee tgaageeetg egtgaagetg 360
accecetgt gegtgaeeet geactgeaee aacetgaaga aegeeaecaa caccaagage 420
agcaactgga aggagatgga cogoggogag atcaagaact gcagottcaa ggtgaccacc 480
agcatccgca acaagatgca gaaggagtac gccctgttct acaagctgga cgtggtgccc 540
atcgacaacg acaacaccag ctacaagctg atcaactgca acaccagcgt gatcacccag 600
gcctgcccca aggtgagctt cgagcccatc cccatccact actgcgcccc cgccggcttc 660
qccatcctqa aqtqcaacqa caaqaaqttc aacggcagcg gcccctgcac caacgtgagc 720
acceptqcaqt qcacccacgg catccgcccc gtggtgagca cccagctgct gctgaacggc 780
ageotggeeg aggagggegt ggtgateege agegagaaet teaeegaeaa egeeaagaee 840
atcatcgtgc agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacacc 900
cgcaagagca tcaccatcgg ccccggccgc gccttctacg ccaccggcga catcatcggc 960
gacatcegee aggeecactg caacatcage ggegagaagt ggaacaacac cetgaageag 1020
atogtgacca agotgcaggo ccagttoggo aacaagacca togtgttcaa gcagagcago 1080
ggcggcgacc ccqaqatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacaqcaccc aqctqttcaa caqcacctqq aacaacacca tcqqccccaa caacaccaac 1200
ggcaccatca cootgecetg cegcateaag cagateggeg geatgtaege ceeecceate 1260
cgcggccaga tccgctgcag cagcaacatc accggcctgc tgctgacccg cgacggcggc 1320
aaggagatca gcaacaccac cgagatcttc cgccccggcg gcggcgacat gcgcgacaac 1380
tggcgcagcg agctgtacaa gtacaaggtg gtgaagatcg agcccctggg cgtggccccc 1440
accaaggcca agegeeget ggtgeagege gagaagegeg eegtgaeeet gggegeeatg 1500
ttcctgggct tcctgggcgc cgccggcagc accatgggcg cccgcagcct gaccctgacc 1560
gtgcaggccc gccagctgct gagcggcatc gtgcagcagc agaacaacct gctgcgcgcc 1620
atcgaggccc agcagcacct gctgcagctg accgtgtggg gcatcaagca gctgcaggcc 1680
egegtgetgg eegtggageg etacetgaag gaceageage tgetgggeat etggggetge 1740
ageggeaage tgatetgeae cacegeegtg ceetggaacg ceagetggag caacaagage 1800
ctggaccaga tctggaacaa catgacctgg atggagtggg agcgcgagat cgacaactac 1860
accaacctga tctacaccct gatcgaggag agccagaacc agcaggagaa gaacgagcag 1920
gagetgetgg agetggaeaa gtgggeeage etgtggaaet ggttegaeat cageaagtgg 1980
```

```
ctgtggtaca tcaagatctt catcatgatc gtgggcggcc tggtgggcct gcgcatcgtg 2040
ttcaccgtgc tgagcatcgt gaaccgcgtg cgccagggct acagccccct gagcttccag 2100
accegettee eegeceeeg eggeeeegae egeceegagg geategagga ggagggegge 2160
gagogogaco gogacogoag cagococotg gtgcacggcc tgctggccct gatctgggac 2220
gacctgcgca gcctgtgcct gttcagctac caccgcctgc gcgacctgat cctgatcgcc 2280
gcccgcatcg tggagctgct gggccgccgc ggctgggagg ccctgaagta ctggggcaac 2340
ctgctgcagt actggatcca ggagctgaag aacagcgccg tgagcctgtt cgacgccatc 2400
gccatcgccg tggccgaggg caccgaccgc atcatcgagg tggcccagcg catcggccgc 2460
gccttcctgc acatccccg ccgcatccgc cagggcttcg agcgcgcct gctgtaactc 2520
<210> 16
<211> 2517
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Gln422-Tyr435
<400> 16
gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccqtqtqqa aqqaqqccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtge acaacgtgtg ggccacccac gcctgcgtge ccaccgacce caacccccag 240
gagategtge tggagaaegt gacegagaae tteaacatgt ggaagaacaa catggtggag 300
cagatgeacg aggacateat cageetgtgg gaccagagee tgaageeetg egtgaagetg 360
accccctgt gcgtgaccct gcactgcacc aacctgaaga acgccaccaa caccaagagc 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcateegea acaagatgea gaaggagtae geeetgttet acaagetgga egtggtgeee 540
ategacaacg acaacaccag ctacaagetg atcaactgca acaccagegt gateacccag 600
gcctqccca aqqtqaqctt cgagcccatc cccatccact actgcgcccc cgccggcttc 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gcccctgcac caacgtgagc 720
acceptgcagt gcacccacgg catccgcccc gtggtgagca cccagctgct gctgaacggc 780
agcctggccg aggaggcgt ggtgatccgc agcgagaact tcaccgacaa cgccaagacc 840
atcatcgtgc agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacacc 900
cgcaagagca tcaccatcgg ccccggccgc gccttctacg ccaccggcga catcatcggc 960
gacatcogcc aggcccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
ategtgacca agetgeagge ceagttegge aacaagacca tegtgtteaa geagageage 1080
ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcaccc agctgttcaa cagcacctgg aacaacacca tcggccccaa caacaccaac 1200
ggcaccatca coetgeeetg eegeateaag eagggegget acgeeeece cateegegge 1260
cagatocget geageageaa catcacegge etgetgetga ceegegaegg eggeaaggag 1320
atcagcaaca ccaccgagat cttccgcccc ggcggcggcg acatgcgcga caactggcgc 1380
agcgagctgt acaagtacaa ggtggtgaag atcgagcccc tgggcgtggc ccccaccaag 1440
gccaagegee gegtggtgea gegegagaag egegeegtga ceetgggege catgtteetg 1500
ggetteetgg gegeegeegg cageaceatg ggegeeegea geetgaeeet gaeegtgeag 1560
gcccgccagc tgctgagcgg catcgtgcag cagcagaaca acctgctgcg cgccatcgag 1620
gcccagcagc acctgctgca gctgaccgtg tggggcatca agcagctgca ggcccgcgtg 1680
ctggccgtgg agcgctacct gaaggaccag cagctgctgg gcatctgggg ctgcagcggc 1740
aagctgatct gcaccacogc cgtgccctgg aacgccagct ggagcaacaa gagcctggac 1800
cagatetgga acaacatgac etggatggag tgggagegeg agategacaa etacaccaac 1860
ctgatctaca ccctgatcga ggagagccag aaccagcagg agaagaacga gcaggagctg 1920
ctggagctgg acaagtgggc cagcctgtgg aactggttcg acatcagcaa gtggctgtgg 1980
tacatcaaga tetteateat gategtggge ggeetggtgg geetgegeat egtgtteace 2040
gtgctgagca tcgtgaaccg cgtgcgccag ggctacagcc ccctgagctt ccagacccgc 2100
ttccccgccc cccgcggccc cgaccgcccc gagggcatcg aggaggggg cggcgagcgc 2160
gaccgcgacc gcagcagccc cctggtgcac ggcctgctgg ccctgatctg ggacgacctg 2220
cgcagcctgt gcctgttcag ctaccaccgc ctgcgcgacc tgatcctgat cgccgcccgc 2280
atcgtggagc tgctgggccg ccgcggctgg gaggccctga agtactgggg caacctgctg 2340
```

```
cagtactgga tecaggaget gaagaacage geegtgagee tgttegaege categeeate 2400
geogtggeog agggeacega eegeateate gaggtggeee agegeategg eegegeette 2460
ctgcacatcc cccgccgcat ccgccagggc ttcgagcgcg ccctgctgta actcgag
<210> 17
<211> 2517
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Gln422-Tyr435B
<400> 17
qaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccqtqtqqa aqqaqqccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accqaqqtqc acaacqtqtq ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagategtge tggagaaegt gaeegagaae tteaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
acceccetgt gegtgaccet geactgeace aacetgaaga acgecaceaa caccaagage 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcateegea acaagatgea gaaggagtae geeetgttet acaagetgga egtggtgeee 540
ategacaacq acaacaccag ctacaagctg atcaactgca acaccagcgt gatcacccag 600
quetquecea aqqtgagett egageceate eccatecaet actgegeece egeeggette 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gcccctgcac caacgtgagc 720
acceptgeagt geacceaegg cateegeece gtggtgagea ceeagetget getgaaegge 780
agectggeeg aggagggegt ggtgateege agegagaact teacegacaa egecaagaee 840
atcatcgtgc agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacacc 900
cgcaagagca tcaccatcgg ccccggccgc gccttctacg ccaccggcga catcatcggc 960
gacatccqcc aggcccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
atcqtqacca agctqcaqqc ccaqttcqqc aacaaqacca tcqtqttcaa qcaqaqcaqc 1080
qqcqqcqacc ccqaqatcqt gatqcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcaccc agctgttcaa cagcacctgg aacaacacca tcggccccaa caacaccaac 1200
ggcaccatca coctgecetg cegeateaag caggeeeeet aegeeeeee cateegegge 1260
cagatocgct gcagcagcaa catcacegge ctgctgctga cccgcgacgg cggcaaggag 1320
atcagcaaca ccaccgagat cttccgcccc ggcggcggcg acatgcgcga caactggcgc 1380
agcgagctgt acaagtacaa ggtggtgaag atcgagcccc tgggcgtggc ccccaccaag 1440
qccaaqcqcc qcqtqqtqca gcgcgagaag cgcgccgtga ccctgggcgc catgttcctg 1500
ggcttcctgg gcgccgcgg cagcaccatg ggcgcccgca gcctgaccct gaccgtgcag 1560
qcccqccaqc tqctqaqcqq catcqtqcaq caqcaqaaca acctqctqcq cqccatcgag 1620
gcccagcagc acctgctgca gctgaccgtg tggggcatca agcagctgca ggcccgcgtg 1680
ctggccgtgg agcgctacct gaaggaccag cagctgctgg gcatctgggg ctgcagcggc 1740
aagetgatet geaccacege egtgeeetgg aacgeeaget ggageaacaa gageetggae 1800
cagatotgga acaacatgac otggatggag tgggagcgcg agatogacaa otacaccaac 1860
ctgatctaca ccctgatcga ggagagccag aaccagcagg agaagaacga gcaggagctg 1920
ctggagctgg acaagtgggc cagcctgtgg aactggttcg acatcagcaa gtggctgtgg 1980
tacatcaaqa tottcatcat gatcgtgggc ggcctggtgg gcctgcgcat cgtgttcacc 2040
gtgctgagca tcgtgaaccg cgtgcgccag ggctacagcc ccctgagctt ccagacccgc 2100
ttccccgccc cccgcggccc cgaccgcccc gagggcatcg aggaggaggg cggcgagcgc 2160
gacegegace geageageee eetggtgeac ggeetgetgg eeetgatetg ggaegacetg 2220
egeageetgt geetgtteag etaceaeege etgegegace tgateetgat egeegeeege 2280
atcgtggagc tgctgggccg ccgcggctgg gaggccctga agtactgggg caacctgctg 2340
cagtactgga tccaggagct gaagaacagc gccgtgagcc tgttcgacgc catcgccatc 2400
geogtggeeg agggeacega cegeateate gaggtggeec agegeategg cegegeette 2460
ctgcacatcc cccgccgcat ccgccagggc ttcgagcgcg ccctgctgta actcgag
<210> 18
<211> 2322
<212> DNA
```

```
<213> Artificial Sequence
<223> Description of Artificial Sequence: Leu122-Ser199;
     Arg426-Gly431
gaattegeea eeatggatge aatgaagaga gggetetget gtgtgetget getgtgtgqa 60
gcagtetteg tttegeccag egeegtggag aagetgtggg tgacegtgta etaeggegtg 120
cccqtqtqqa aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accqaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgeacg aggacateat cageetgtgg gaccagagee tgaageeetg egtgaagetg 360
ggcaacagcg tgatcaccca ggcctgcccc aaggtgagct tcgagcccat ccccatccac 420
tactqcqccc ccqccqqctt cqccatcctg aagtgcaacg acaagaagtt caacggcagc 480
ggcccctgca ccaacgtgag caccgtgcag tgcacccacg gcatccgccc cgtggtgagc 540
acccagetge tgctgaacgg cageetggee gaggagggeg tggtgateeg cagegagaac 600
ttcaccqaca acqccaaqac catcatcqtq cagctgaagg agagcgtgga gatcaactgc 660
accequeca acaacaacac eegcaagage atcaccateg geeeeggeeg egeettetac 720
gccaccggcg acatcatcgg cgacatccgc caggcccact gcaacatcag cggcgagaag 780
tggaacaaca ccctgaagca gatcgtgacc aagctgcagg cccagttcgg caacaagacc 840
atcgtgttca agcagagcag cggcggcgac cccgagatcg tgatgcacag cttcaactgc 900
ggcggcgagt tettetactg caacagcace cagetgttea acagcacetg gaacaacace 960
ateggeecca acaacaccaa eggeaccate accetgeect geegeateaa geagateate 1020
aaccgcggcg gcggcaaggc catgtacgcc cccccatcc gcggccagat ccgctgcagc 1080
aqcaacatca coqqcctqct qctgacccgc gacggcggca aggagatcag caacaccacc 1140
gagatettee geeeeggegg eggegacatg egegacaact ggegeagega getgtacaag 1200
tacaaggtgg tgaagatcga gcccctgggc gtggccccca ccaaggccaa gcgccgcgtg 1260
qtqcaqcqcq aqaaqcqcqc cqttgaccctq ggcgccatgt tcctgggctt cctgggcgcc 1320
geeggeagea ceatgggege eegcageetg accetgaceg tgeaggeeeg ceagetgetg 1380
ageggeateg tgeageagea gaacaacetg etgegegeea tegaggeeea geageacetg 1440
ctgcagctga ccgtgtgggg catcaagcag ctgcaggccc gcgtgctggc cgtggagcgc 1500
tacctgaagg accagcagct gctgggcatc tggggctgca gcggcaagct gatctgcacc 1560
accqccqtqc cctqqaacqc caqctqqaqc aacaagagcc tggaccagat ctggaacaac 1620
atgacetgga tggagtggga gegegagate gacaactaca ceaacetgat etacaceetg 1680
atcqaqqaqa qccaqaacca qcaqqaqaaq aacgaqcagg aqctqctgqa gctggacaag 1740
tqqqccaqcc tqtqqaactq qttcgacatc agcaagtggc tgtggtacat caagatcttc 1800
atcatgatcg tgggcggcct ggtgggcctg cgcatcgtgt tcaccgtgct gagcatcgtg 1860
aaccgcgtgc gccagggcta cagcccctg agcttccaga cccgcttccc cgcccccgc 1920
ggccccgacc gccccgaggg catcgaggag gagggcggcg agcgcgaccg cgaccgcagc 1980
ageccetqq tqcacqqcct gctggccctg atctgggacg acctgcgcag cctgtgcctg 2040
ttcagctacc accycctgcg cqacctgatc ctgatcgccg cccgcatcgt ggagctgctg 2100
qqccqccqcq qctqqqaqqc cctqaagtac tggggcaacc tgctgcagta ctggatccaq 2160
qaqctqaaqa acaqcqccqt gagcctgttc gacgccatcg ccatcgccgt ggccgagggc 2220
accgaccgca tcatcgaggt ggcccagcgc atcggccgcg ccttcctgca catcccccgc 2280
cgcatccgcc agggcttcga gcgcgccctg ctgtaactcg ag
                                                                  2322
<210> 19
<211> 2322
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Leu122-Ser199;
      Arg426-Lys432
<400> 19
qaatteqeea ccatqqatqe aatqaaqaqa qqqctctqct qtqtqctqct qctqttqqa 60
gcaqtcttcq tttcqcccaq cqccqtqqag aagctgtggg tqaccqtqta ctacqqcqtq 120
```

```
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
ggcaacagcg tgatcaccca ggcctgcccc aaggtgagct tcgagcccat ccccatccac 420
tactgcgccc ccgccggctt cgccatcctg aagtgcaacg acaagaagtt caacggcagc 480
ggcccctgca ccaacgtgag caccgtgcag tgcacccacg gcatccgccc cgtggtgagc 540
acccagetge tgetgaacgg cageetggee gaggagggeg tggtgateeg cagegagaac 600
ttcaccgaca acgccaagac catcatcgtg cagctgaagg agagcgtgga gatcaactgc 660
accegececa acaacaacac eegcaagage atcaceateg geeceggeeg egcettetae 720
gccaccggcg acatcatcgg cgacatccgc caggcccact gcaacatcag cggcgagaag 780
tggaacaaca ccctgaagca gatcgtgacc aagctgcagg cccagttcgg caacaagacc 840
atogtgttca agcagagcag oggoggogac occgagatog tgatgcacag ottcaactgc 900
ggcggcgagt tettetactg caacagcace cagetgttea acageacetg gaacaacace 960
ateggececa acaacaceaa eggeaceate accetgecet geegeateaa geagateate 1020
aaccgcggcg gcaacaaggc catgtacgcc cccccatcc gcggccagat ccgctgcagc 1080
agcaacatca ceggeetget getgaeeege gaeggeggea aggagateag caacaccace 1140
gagatettee geeeggegg eggegacatg egegacaact ggegeagega getgtacaag 1200
tacaaggtgg tgaagatcga gcccctgggc gtggccccca ccaaggccaa gcgccgcgtg 1260
gtgcagcgcg agaagcgcgc cgtgaccctg ggcgccatgt tcctgggctt cctgggcgcc 1320
geoggeagea ceatgggege eegeageetg accetgaceg tgeaggeeeg ceagetgetg 1380
ageggeateg tgeageagea gaacaacetg etgegegeea tegaggeeea geageacetg 1440
ctgcagctga ccgtgtgggg catcaagcag ctgcaggccc gcgtgctggc cgtggagcgc 1500
tacctgaagg accagcagct gctgggcatc tggggctgca gcggcaagct gatctgcacc 1560
accgccgtgc cctggaacgc cagctggagc aacaagagcc tggaccagat ctggaacaac 1620
atgacctgga tggagtggga gcgcgagatc gacaactaca ccaacctgat ctacaccctg 1680
atcgaggaga gccagaacca gcaggagaag aacgagcagg agctgctgga gctggacaag 1740
tgggccagcc tgtggaactg gttcgacatc agcaagtggc tgtggtacat caagatcttc 1800
atcatgateg tgggeggeet ggtgggeetg egeategtgt teacegtget gageategtg 1860
aaccgcgtgc gccagggcta cagcccctg agcttccaga cccgcttccc cgcccccgc 1920
ggccccgacc gccccgaggg catcgaggag gagggcggcg agcgcgaccg cgaccgcagc 1980
agccccctgg tgcacggcct gctggccctg atctgggacg acctgcgcag cctgtgcctg 2040
ttcagctacc accgcctgcg cgacctgatc ctgatcgccg cccgcatcgt ggagctgctg 2100
ggccgccgcg gctgggaggc cctgaagtac tggggcaacc tgctgcagta ctggatccag 2160
gagetgaaga acagegeegt gageetgtte gaegeeateg ceategeegt ggeegaggge 2220
accgaccgca tcatcgaggt ggcccagcgc atcggccgcg ccttcctgca catcccccgc 2280
cgcatccgcc agggettcga gcgcgccctg ctgtaactcg ag
                                                                  2322
<210> 20
<211> 2322
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Leu122-Ser199;
      Trp427-Gly431
<400> 20
gaattegeea ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
ggcaacagcg tgatcaccca ggcctgcccc aaggtgagct tcgagcccat ccccatccac 420
tactgegece eegeeggett egecateetg aagtgeaaeg acaagaagtt caaeggeage 480
ggcccctgca ccaacgtgag caccgtgcag tgcacccacg gcatccgccc cgtggtgagc 540
acccagetge tgetgaaegg cageetggee gaggagggeg tggtgateeg cagegagaae 600
ttcaccgaca acgccaagac catcatcgtg cagctgaagg agagcgtgga gatcaactgc 660
```

```
accegeeeca acaacaacac eegcaagage atcaccateg geeeeggeeg egeettetae 720
gecaceggeg acateategg egacateege caggeccact geaacateag eggegagaag 780
tggaacaaca ccctgaagca gatcgtgacc aagctgcagg cccagttcgg caacaagacc 840
atogtqttca agcagagcag cggcggcgac cccgagatcg tgatgcacag cttcaactgc 900
ggcggcgagt tettetactg caacagcace cagetgttea acagcacetg gaacaacace 960
ateggececa acaacaccaa eggeaccate accetgecet geegeateaa geagateate 1020
aaccqctqqq qcqqcaaqqc catqtacqcc cccccatcc qcqqccagat ccqctqcagc 1080
agcaacatca coggoctgot gotgaccogo gacggoggca aggagatcag caacaccacc 1140
gagatettee geeeggegg eggegacatg egegacaaet ggegeagega getgtacaag 1200
tacaaqqtqq tgaaqatcqa gcccctgggc gtggcccca ccaaggccaa gcgccgcgtg 1260
gtgcagcgcg agaagcgcgc cgtgaccctg ggcgccatgt tcctgggctt cctgggcgcc 1320
geoggeagea ceatgggege eegeageetg accetgaceg tgeaggeeeg ceagetgetg 1380
ageggeateg tgeageagea gaacaacetg etgegegeea tegaggeeea geageacetg 1440
ctgcagctga ccgtgtgggg catcaagcag ctgcaggccc gcgtgctggc cgtggagcgc 1500
tacctgaagg accagcagct gctgggcatc tggggctgca gcggcaagct gatctgcacc 1560
accqccqtqc cctqqaacqc cagctggagc aacaagagcc tggaccagat ctggaacaac 1620
atgacctgga tggagtggga gcgcgagatc gacaactaca ccaacctgat ctacacctg 1680
atcgaggaga gccagaacca gcaggagaag aacgagcagg agctgctgga gctggacaag 1740
tgggccagcc tgtggaactg gttcgacatc agcaagtggc tgtggtacat caagatcttc 1800
atcatgatcg tgggcggcct ggtgggcctg cgcatcgtgt tcaccgtgct gagcatcgtg 1860
aaccgcgtgc gccagggcta cagccccctg agcttccaga cccgcttccc cgccccccgc 1920
ggccccgacc gccccgaggg catcgaggag gagggcggcg agcgcgaccg cgaccgcagc 1980
ageoccetqq tqcacqqcct getggccetg atctgggacg acctgcgcag cctgtgcctg 2040
ttcagctacc accectege cgacctgatc ctgatcgccg cccgcatcgt ggagctgctg 2100
ggccgccgcg gctggggggc cctgaagtac tggggcaacc tgctgcagta ctggatccag 2160
gagetgaaga acagegeegt gageetgtte gaegeeateg ceategeegt ggeegaggge 2220
accgaccgca tcatcgaggt ggcccagcgc atcggccgcg ccttcctgca catccccgc 2280
cgcatccgcc agggcttcga gcgcgccctg ctgtaactcg ag
                                                                  2322
<210> 21
<211> 2310
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Lys121-Val200;
     Asn425-Lys432
<400> 21
gaatteqeea eeatqqatqe aatgaagaga gggetetget gtgtgctget getgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagategtge tggagaaegt gaeegagaae tteaacatgt ggaagaacaa catggtggag 300
cagatgeacg aggacateat cageetgtgg gaccagagee tgaageeetg egtgaaggee 360
cccgtgatca cccaggcctg ccccaaggtg agcttcgagc ccatccccat ccactactgc 420
gcccccgccg gcttcgccat cctgaagtgc aacgacaaga agttcaacgg cagcggcccc 480
tgcaccaacg tgagcaccgt gcagtgcacc cacggcatcc gccccgtggt gagcacccag 540
ctgctgctga acggcagcct ggccgaggag ggcgtggtga tccgcagcga gaacttcacc 600
gacaacgcca agaccatcat cgtgcagctg aaggagagcg tggagatcaa ctgcaccegc 660
cccaacaaca acaccegcaa gagcatcacc atcggccccg gccgcgctt ctacgccacc 720
ggcgacatca tcggcgacat ccgccaggcc cactgcaaca tcagcggcga gaagtggaac 780
aacaccctga agcagatcgt gaccaagctg caggcccagt tcggcaacaa gaccatcgtg 840
ttcaaqcaqa qcaqcqqqqq cgaccccgag atcgtqatqc acaqcttcaa ctqcqqcqqc 900
gagttettet actqcaacag caccagetg ttcaacagea cetqqaacaa caccategge 960
cccaacaaca ccaacggcac catcaccctg ccctgccgca tcaagcagat catcaacgcc 1020
cccaaggeca tgtacgecc ccccateege ggccagatee getgeageag caacateace 1080
ggcctgctgc tqacccgcga cggcggcaag gagatcagca acaccaccga gatcttccgc 1140
cccggcggcg gcgacatgcg cgacaactgg cgcagcgagc tgtacaagta caaggtggtg 1200
```

```
aagategage ceetgggegt ggeececace aaggeeaage geegegtggt geagegegag 1260
aagegegeeg tgaccetggg egecatgtte etgggettee tgggegeege eggeageace 1320
atgggcgccc gcagcctgac cctgaccgtg caggcccgcc agctgctgag cggcatcgtg 1380
cagcagcaga acaacctgct gegegecate gaggeceage ageacctget geagetgace 1440
gtgtggggca tcaagcagct gcaggcccgc gtgctggccg tggagcgcta cctgaaggac 1500
cagcagetge tgggcatetg gggctgcage ggcaagetga tetgcaceae egeegtgeee 1560
tggaacgcca gctggagcaa caagagcctg gaccagatct ggaacaacat gacctggatg 1620
gagtgggagc gcgagatcga caactacacc aacctgatct acaccctgat cgaggagagc 1680
cagaaccagc aggagaagaa cgagcaggag ctgctggagc tggacaagtg ggccagcctg 1740
tggaactggt tcgacatcag caagtggctg tggtacatca agatcttcat catgatcgtg 1800
ggcggcctgg tgggcctgcg catcgtgttc accgtgctga gcatcgtgaa ccgcgtgcgc 1860
cagggetaca geoccetgag ettecagace egetteeceg ecceegegg eccegacege 1920
cccgagggca tcgaggagga gggcggcgag cgcgaccgcg accgcagcag ccccctggtg 1980
cacggcctgc tggccctgat ctgggacgac ctgcgcagcc tgtgcctgtt cagctaccac 2040
egectgegeg acetgatect gategeegee egeategtgg agetgetggg eegeegegge 2100
tgggaggccc tgaagtactg gggcaacctg ctgcagtact ggatccagga gctgaagaac 2160
agegeegtga geetgttega egeeategee ategeegtgg eegagggeae egaeegeate 2220
atcgaggtgg cccagcgcat cggccgcgcc ttcctgcaca tcccccgccg catccgccag 2280
ggcttcgagc gcgccctgct gtaactcgag
<210> 22
<211> 2298
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Val120-Ile201;
      Ile424-Ala433
<400> 22
gaattegeea ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagategtge tggagaaegt gaeegagaae tteaacatgt ggaagaaeaa catggtggag 300
cagatgcacq aggacatcat cagcetgtgg gaccagagec tgaagcectg cgtgggcggc 360
atcacccagg cetgececaa ggtgagette gageceatee ceatecacta etgegeecee 420
geoggetteg ceatectgaa gtgcaacgac aagaagttea acggcagegg cecetgeace 480
aacgtgagca ccgtgcagtg cacccacggc atccgccccg tggtgagcac ccagctgctg 540
ctgaacggca gcctggccga ggagggcgtg gtgatccgca gcgagaactt caccgacaac 600
gccaagacca tcatcgtgca gctgaaggag agcgtggaga tcaactgcac ccgccccaac 660
aacaacacc gcaagagcat caccategge eceggeegeg cettetaege caceggegae 720
atcateggeg acateegeea ggeecaetge aacateageg gegagaagtg gaacaacace 780
ctqaaqcaqa tcqtqaccaa qctgcaggcc cagttcggca acaagaccat cgtgttcaag 840
cagageageg geggegaeec egagategtg atgeacaget teaactgegg eggegagtte 900
ttotactgca acagcaccca gotgttcaac agcacctgga acaacaccat oggcoccaac 960
aacaccaacg gcaccatcac cctgccctgc cgcatcaagc agatcatcgg cggcgccatg 1020
tacgcccccc ccatccgcgg ccagatccgc tgcagcagca acatcaccgg cctgctgctg 1080
accegegaeg geggeaagga gateageaac accaeegaga tetteegeee eggeggegge 1140
gacatgegeg acaactggeg cagegagetg tacaagtaca aggtggtgaa gategageec 1200
etgggegtgg ecceaceaa ggecaagege egegtggtge agegegagaa gegegeegtg 1260
accetqqqcq ccatqttcct gggcttcctg ggcgccgccg gcagcaccat gggcgcccgc 1320
agcetgacce tgacegtgea ggeecgeeag etgetgageg geategtgea geageagaae 1380
aacctgctgc gcgccatcga ggcccagcag cacctgctgc agctgaccgt gtggggcatc 1440
aagcagctgc aggcccgcgt gctggccgtg gagcgctacc tgaaggacca gcagctgctg 1500
ggcatctggg gctgcagcgg caagctgatc tgcaccaccg ccgtgccctg gaacgccagc 1560
tggagcaaca agagcctgga ccagatctgg aacaacatga cctggatgga gtgggagcgc 1620
gagategaca actacaceaa cetgatetac accetgateg aggagageca gaaceageag 1680
gagaagaacg agcaggagct gctggagctg gacaagtggg ccagcctgtg gaactggttc 1740
```

```
gacatcagca agtggctgtg gtacatcaag atcttcatca tgatcgtggg cggcctggtg 1800
ggcctgcgca tcgtgttcac cgtgctgagc atcgtgaacc gcgtgcgcca gggctacagc 1860
cccctgaget tecagacccg cttccccgcc ccccgcggcc ccgaccgccc cgagggcatc 1920
gaggaggagg geggegageg egacegegae egeageagee eeetggtgea eggeetgetg 1980
gccctgatct gggacgacct gcgcagcctg tgcctgttca gctaccaccg cctgcgcgac 2040
ctgatcctga tegecgeeg catcgtggag ctgctgggee geegeggetg ggaggeeetg 2100
aagtactggg gcaacctgct gcagtactgg atccaggagc tgaagaacag cgccgtgagc 2160
etgttegaeg ceategeeat egeegtggee gagggeaceg acegeateat egaggtggee 2220
cagogoatog geogogoett cetgoacate eccegoegoa teegoeaggg ettegagege 2280
gccctgctgt aactcgag
<210> 23
<211> 2298
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence:
      Val120-Ile201B; Ile424-Ala433
<400> 23
gaattegeea eeatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
geagtetteg tttegeecag egeegtggag aagetgtggg tgacegtgta etaeggegtg 120
cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagategtge tggagaaegt gacegagaae tteaacatgt ggaagaacaa catggtggag 300
cagatgeacg aggacateat cageetgtgg gaccagagee tgaageeetg egtgeeegge 360
atcacccagg cctgccccaa ggtgagcttc gagcccatcc ccatccacta ctgcgccccc 420
geoggetteg ceateetgaa gtgcaacgac aagaagttea aeggeagegg cecetgeace 480
aacgtgagca ccgtgcagtg cacccacggc atccgccccg tggtgagcac ccagctgctg 540
ctgaacggca gcctggccga ggagggcgtg gtgatccgca gcgagaactt caccgacaac 600
gccaagacca tcatcgtgca gctgaaggag agcgtggaga tcaactgcac ccgccccaac 660
aacaacacce gcaagagcat caccategge eceggeegeg cettetacge caceggegae 720
atcateggeg acateegeea ggeccaetge aacateageg gegagaagtg gaacaacace 780
ctqaaqcaqa tcqtqaccaa qctqcaqqcc cagttcqqca acaaqaccat cgtqttcaag 840
cagageageq geggegaccc egagategtg atgeacaget teaactgegg eggegagtte 900
ttctactgca acagcaccca gctgttcaac agcacctgga acaacaccat cggccccaac 960
aacaccaacg gcaccatcac cetgecetge egeatcaage agateategg eggegecatg 1020
tacgccccc ccatccgcgg ccagatccgc tgcagcagca acatcaccgg cctgctgctg 1080
accegegacg geggeaagga gateageaac accaeegaga tetteegeee eggeggegge 1140
gacatgcgcg acaactggcg cagcgagctg tacaagtaca aggtggtgaa gatcgagccc 1200
ctgggcgtgg ccccaccaa ggccaagcgc cgcgtggtgc agcgcgagaa gcgcgccgtg 1260
accotgggeg ceatgitest gggetteetg ggegeegeeg geageaceat gggegeege 1320
agectqaeec tqaeeqtqea ggeeegeeag etgetgageg geategtgea geaqeagaac 1380
aacctgctgc gcgccatcga ggcccagcag cacctgctgc agctgaccgt gtggggcatc 1440
aagcagctgc aggcccgcgt gctggccgtg gagcgctacc tgaaggacca gcagctgctg 1500
ggcatctggg gctgcagcgg caagctgatc tgcaccaccg ccgtgccctg gaacgccagc 1560
tggagcaaca agagcctgga ccagatctgg aacaacatga cctggatgga gtgggagcgc 1620
gagategaca actacaceaa cetgatetae accetgateg aggagageea gaaceageag 1680
gagaagaacg agcaggagct gctggagctg gacaagtggg ccagcctgtg gaactggttc 1740
gacatcagca agtggctgtg gtacatcaag atcttcatca tgatcgtggg cggcctggtg 1800
ggcctgcqca tcqtqttcac cgtgctgagc atcgtgaacc gcgtgcgcca gggctacagc 1860
cocctgaget tecagacceg etteccegee eccegeggee ecgacegeee egagggeate 1920
gaggaggagg geggegageg egacegegae egeageagee ecetggtgea eggeetgetg 1980
geoetgatet gggacgaeet gegeageetg tgeetgttea getaceaeeg eetgegegae 2040
ctgatectga tegeogeceg categtggag etgetgggee geogeggetg ggaggecetg 2100
aagtactggg gcaacctgct gcagtactgg atccaggagc tgaagaacag cgccgtgagc 2160
ctgttcgacg ccatcgccat cgccgtggcc gagggcaccg accgcatcat cgaggtggcc 2220
cagogoateq qeeqecett cetqeacate ceeeqeeqea teeqeeaqqq etteqaqeqe 2280
```

```
gccctgctgt aactcgag
                                                                  2298
<210> 24
<211> 2298
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Val120-Thr202;
      Ile424-Ala433
<400> 24
gaattegeca ecatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccqtqtqqa aqqaqqcac caccacctq ttctqcqcca qcqacqccaa qqcctacqac 180
accqagqtqc acaacqtqtq ggccacccac gcctgcgtgc ccaccqaccc caacccccaq 240
gagategtge tggagaaegt gaeegagaae tteaacatgt ggaagaacaa catqqtqqaq 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgggeggc 360
gccacccagg cctgccccaa ggtgagcttc gagcccatcc ccatccacta ctqcqcccc 420
geoggetteg ccatectgaa gtgcaacgac aagaagttea acggcagegg cccctgcacc 480
aacqtqaqca ccqtqcaqtq cacccacggc atccgccccq tqqtqaqcac ccaqctqctq 540
ctgaacggca gcctggccga ggagggcgtg gtgatccgca gcgagaactt caccgacaac 600
gccaagacca tcatcgtgca gctgaaggag agcgtggaga tcaactgcac ccgccccaac 660
aacaacacc gcaagagcat caccategge ceeggeegeg cettetaege caceggeegae 720
atcateggeg acateegeea ggeccaetge aacateageg gegagaagtg gaacaacace 780
ctgaagcaga tcgtgaccaa gctgcaggcc cagttcggca acaagaccat cgtgttcaag 840
cagageageg geggegacce egagategtg atgeacaget teaactgegg eggegagtte 900
ttctactgca acagcaccca gctgttcaac agcacctgga acaacaccat cggccccaac 960
aacaccaacg gcaccatcac cctgccctgc cgcatcaagc agatcatcgg cggcgccatg 1020
tacqccccc ccatccqcqq ccaqatccqc tqcaqcaqca acatcaccqq cctqctqctq 1080
accegegacg geggeaagga gateageaac accacegaga tetteegeec eggeggegge 1140
qacatqqqqq acaactqqcq caqcqaqctq tacaaqtaca aqqtqqtqaa qatcqaqccc 1200
ctgggcgtgg ccccaccaa ggccaagcgc cgcgtggtgc agcgcgagaa gcgcgccgtg 1260
accetgggeg coatgitect gggetteetg ggegeegeeg geageaceat gggegeeege 1320
agectgaece tqaccqtqca ggcccgccag ctgctgagcg gcatcqtqca gcaqcaqaac 1380
aacetgetge gegeeatega ggeecagcag cacetgetge agetgacegt gtggggeate 1440
aagcagctgc aggcccqcqt gctggccgtg gagcgctacc tgaaggacca gcagctgctg 1500
ggcatctggg gctgcagcgg caagctgatc tgcaccaccg ccgtgccctg gaacgccagc 1560
tggagcaaca agagcctgga ccagatctgg aacaacatga cctggatgga gtgggagcgc 1620
qaqatcqaca actacaccaa cctqatctac accctqatcq aqqaqaqcca qaaccaqcaq 1680
gagaaqaacq aqcaqqaqct qctggagctg gacaagtggg ccaqcctgtg gaactqqttc 1740
gacatcagca agtggctgtg gtacatcaag atcttcatca tgatcgtggg cggcctggtg 1800
ggcctqcqca tcqtqttcac cgtgctgagc atcgtgaacc gcgtqcqcca qqqctacaqc 1860
cccctgaget tecagacceg etteccegec eccegegec ecgacegece egagggeate 1920
gaggaggagg geggegageg egacegegae egeageagee eeetggtgea eggeetgetg 1980
geoctgatet gggacgaeet gegeageetg tgeetgttea getaecaeeg eetgegegae 2040
ctgatcctga tcgccgcccg catcgtggag ctgctgggcc gccgcggctg ggaggccctg 2100
aagtactqqq qcaacctqct gcagtactgg atccaggagc tqaaqaacag cqccqtqaqc 2160
ctgttcgacg ccatcgccat cgccgtggcc gagggcaccg accgcatcat cgaggtggcc 2220
cagegeateg geogegeett cetgeacate eccegeegea teegeeaggg ettegagege 2280
gccctgctgt aactcgag
                                                                  2298
<210> 25
<211> 2358
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Val127-Asn195
```

```
<400> 25
gaattegeea eeatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtetteg tttegcccag egeegtggag aagetgtggg tgacegtgta etaeggegtg 120
cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
acceccetgt gegtggggge agggaactge aacaccageg tgatcaccca ggeetgeece 420
aaggtgaget tegageceat ecceatecae tactgegeee eegeeggett egecateetg 480
aagtgcaacg acaagaagtt caacggcagc ggcccctgca ccaacgtgag caccgtgcag 540
tgcacccacg gcatccgccc cgtggtgagc acccagctgc tgctgaacgg cagcctggcc 600
gaggagggcg tggtgatccg cagcgagaac ttcaccgaca acgccaagac catcatcgtg 660
cagctgaagg agagcgtgga gatcaactgc accegeeeca acaacaacae eegcaagage 720
atcaccateg geoeeggeeg egeettetae gecaceggeg acateategg egacateege 780
caggeceact geaacateag eggegagaag tggaacaaca eeetgaagea gategtgaee 840
eccgagateg tgatgeacag etteaactge ggeggegagt tettetactg caacageace 960
cagctgttca acagcacctg gaacaacacc atcggcccca acaacaccaa cggcaccatc 1020
accetgeeet geegeateaa geagateate aacegetgge aggaggtggg caaggeeatg 1080
tacgcccccc ccatccgcgg ccagatccgc tgcagcagca acatcaccgg cctgctqctq 1140
accegegacg geggeaagga gateageaac accaeegaga tetteegeee eggeggegge 1200
gacatgcgcg acaactggcg cagcgagctg tacaagtaca aggtggtgaa gatcgagccc 1260
etgggegtgg cccccaccaa ggccaagegc cgcgtggtgc agegegagaa gcgcgccqtq 1320
accotgggcg coatgitect gggetteetg ggcgccgcc gcagcaccat gggcgcccgc 1380
agectgacce tgaccgtgca ggcccgccag ctgctgagcg gcatcgtgca gcagcagaac 1440
aacetgetge gegeeatega ggeeeageag cacetgetge agetgaeegt gtggggeate 1500
aagcagctgc aggcccgcgt gctggccgtg gagcgctacc tgaaggacca gcagctgctg 1560
ggcatctggg gctgcagcgg caagctgatc tgcaccaccg ccgtgccctg gaacgccagc 1620
tggagcaaca agagcctgga ccagatctgg aacaacatga cctggatgga gtgggagcgc 1680
gagategaca actacaceaa cetgatetac accetgateg aggagageca gaaceageag 1740
gagaagaacg agcaggagct gctggagctg gacaagtggg ccagcctgtg gaactggttc 1800
gacatcagca agtggetgtg gtacatcaag atcttcatca tgatcgtggg cggcctggtg 1860
ggcctgcgca tcgtgttcac cgtgctgagc atcgtgaacc gcgtgcgcca gggctacagc 1920
cccctgagct tccagacccg cttccccgcc ccccgcggcc ccgaccgccc cgagggcatc 1980
gaggaggagg geggegageg egacegegae egeageagee eeetggtgea eggeetgetg 2040
gccctgatct gggacgacct gcgcagcctg tgcctgttca gctaccaccg cctgcgcgac 2100
ctgatectga tegeogeceg categtggag etgetgggee geogeggetg ggaggeeetg 2160
aagtactggg gcaacctgct gcagtactgg atccaggagc tgaagaacag cgccgtgagc 2220
ctgttcgacg ccatcgccat cgccgtggcc gagggcaccg accgcatcat cgaggtggcc 2280
cagegeateg geogegeett cetgeacate eccegeegea teegeeaggg ettegagege 2340
gccctgctgt aactcgag
                                                                 2358
<210> 26
<211> 2352
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Val127-Asn195;
      Arg426-Gly431
<400> 26
gaattegeca ccatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccaccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
acceccetgt gegtgggggc agggaactgc aacaccageg tgatcaccca ggcetgeecc 420
```

aaggtgagct	tcgagcccat	ccccatccac	tactgcgccc	ccgccggctt	cgccatcctg	480
aagtgcaacg	acaagaagtt	caacggcagc	ggcccctgca	ccaacgtgag	caccgtgcag	540
tgcacccacg	gcatccgccc	cgtggtgagc	acccagctgc	tgctgaacgg	cagcctggcc	600
gaggagggcg	tggtgatccg	cagcgagaac	ttcaccgaca	acgccaagac	catcatcgtg	660
cagctgaagg	agagcgtgga	gatcaactgc	acccgcccca	acaacaacac	ccgcaagagc	720
atcaccatcg	gccccggccg	cgccttctac	gccaccggcg	acatcatcgg	cgacatccgc	780
caggcccact	gcaacatcag	cggcgagaag	tggaacaaca	ccctgaagca	gatcgtgacc	840
aagctgcagg	cccagttcgg	caacaagacc	atcgtgttca	agcagagcag	cggcggcgac	900
cccgagatcg	tgatgcacag	cttcaactgc	ggcggcgagt	tcttctactg	caacagcacc	960
cagctgttca	acagcacctg	gaacaacacc	atcggcccca	acaacaccaa	cggcaccatc	1020
accctgccct	gccgcatcaa	gcagatcatc	aaccgcggcg	gcggcaaggc	catgtacgcc	1080
cccccatcc	gcggccagat	ccgctgcagc	agcaacatca	ccggcctgct	gctgacccgc	1140
gacggcggca	aggagatcag	caacaccacc	gagatcttcc	gccccggggg	cggcgacatg	1200
cgcgacaact	ggcgcagcga	gctgtacaag	tacaaggtgg	tgaagatcga	gcccctgggc	1260
gtggccccca	ccaaggccaa	gcgccgcgtg	gtgcagcgcg	agaagcgcgc	cgtgaccctg	1320
ggcgccatgt	tcctgggctt	cctgggcgcc	gccggcagca	ccatgggcgc	ccgcagcctg	1380
accctgaccg	tgcaggcccg	ccagctgctg	agcggcatcg	tgcagcagca	gaacaacctg	1440
ctgcgcgcca	tcgaggccca	gcagcacctg	ctgcagctga	ccgtgtgggg	catcaagcag	1500
ctgcaggccc	gcgtgctggc	cgtggagcgc	tacctgaagg	accagcagct	gctgggcatc	1560
tggggctgca	gcggcaagct	gatctgcacc	accgccgtgc	cctggaacgc	cagctggagc	1620
aacaagagcc	tggaccagat	ctggaacaac	atgacctgga	tggagtggga	gcgcgagatc	1680
gacaactaca	ccaacctgat	ctacaccctg	atcgaggaga	gccagaacca	gcaggagaag	1740
aacgagcagg	agctgctgga	gctggacaag	tgggccagcc	tgtggaactg	gttcgacatc	1800
agcaagtggc	tgtggtacat	caagatcttc	atcatgatcg	tgggcggcct	ggtgggcctg	1860
cgcatcgtgt	tcaccgtgct	gagcatcgtg	aaccgcgtgc	gccagggcta	cagccccctg	1920
agcttccaga	cccgcttccc	cgcccccgc	ggccccgacc	gccccgaggg	catcgaggag	1980
gagggcggcg	agcgcgaccg	cgaccgcagc	agccccctgg	tgcacggcct	gctggccctg	2040
atctgggacg	acctgcgcag	cctgtgcctg	ttcagctacc	accgcctgcg	cgacctgatc	2100
ctgatcgccg	cccgcatcgt	ggagctgctg	ggccgccgcg	gctgggaggc	cctgaagtac	2160
tggggcaacc	tgctgcagta	ctggatccag	gagctgaaga	acagcgccgt	gagcctgttc	2220
gacgccatcg	ccatcgccgt	ggccgagggc	accgaccgca	tcatcgaggt	ggcccagcgc	2280
atcggccgcg	ccttcctgca	catcccccgc	cgcatccgcc	agggcttcga	gcgcgccctg	2340
ctgtaactcg	ag					2352